

Dissertation zur Erlangung des Doktorgrades
der Fakultät für Chemie und Pharmazie
der Ludwig-Maximilians-Universität München

Characterization of She2p-dependent mRNP assembly in *Saccharomyces cerevisiae*



Marisa Müller
aus
Pforzheim

2009

Erklärung

Diese Dissertation wurde im Sinne von § 13 Abs. 3 der Promotionsordnung vom 29. Januar 1998 von Herrn Prof. Dr. Patrick Cramer betreut.

Ehrenwörtliche Versicherung

Diese Dissertation wurde selbständig, ohne unerlaubte Hilfe erarbeitet.

München, am 26. Mai 2009

.....

Marisa Müller

Dissertation eingereicht am 26. Mai 2009

1. Gutachter: Prof. Dr. Patrick Cramer
2. Gutachter: Prof. Dr. Klaus Förstemann

Mündliche Prüfung am 8. Juli 2009

Table of Contents

SUMMARY	1
1 INTRODUCTION	2
1.1 mRNA localization – a common mechanism for targeting proteins.....	3
1.2 Assembly and translocation of an mRNA-transport particle.....	4
1.2.1 Assembly of the mRNP (Step1 in Figure 2).....	5
1.2.2 Motor-driven translocation of a translationally silenced mRNP along cytoskeletal tracks (Step 2 in Figure 2)	7
1.2.3 Anchoring of the mRNP and translation of localized mRNAs (Steps 3 and 4 in Figure 2)	9
1.3 mRNA localization in somatic cells.....	9
1.3.1 mRNA localization in fibroblasts	9
1.3.2 mRNA localization in the neuronal system	10
1.4 mRNA localization in oocytes and developing embryos.....	11
1.4.1 mRNA localization in <i>Xenopus laevis</i> oocytes.....	11
1.4.2 Localization of <i>gurken</i> , <i>bicoid</i> and <i>oskar</i> mRNA during <i>Drosophila</i> oogenesis	12
1.5 mRNA localization in <i>Saccharomyces cerevisiae</i>	13
1.5.1 “Mating type switching” – the biological function of <i>ASH1</i> -mRNA localization	14
1.5.2 <i>ASH1</i> and at least 23 additional bud-localized mRNAs.....	15
1.5.3 She2p – the cargo-binding protein.....	17
1.5.4 She3p links the She2p:mRNA complex to the myosin-motor protein Myo4p	19
1.5.5 Additional <i>trans</i> -acting factors involved in <i>ASH1</i> -mRNA localization	20
1.5.6 Current model of <i>ASH1</i> -mRNA localization	22
1.6 Objectives	23
2 RESULTS	24
2.1 Structural studies on the She2p:RNA interaction	24
2.1.1 Purification of She2p variants	24
2.1.2 Designing She2p variants for crystallization	25
2.1.3 Identification of suitable <i>ASH1</i> -RNA fragments for co-crystallization	26
2.1.3.1 She2p-(6-246)- Δ Cys is suitable for crystallization with <i>ASH1</i> -E3 RNA and minimized <i>ASH1</i> zipcodes.....	27
2.1.3.2 Binding studies with She2p- Δ loop and She2p- Δ helixE.....	28
2.1.4 Co-crystallization experiments with <i>ASH1</i> -E3, E1min, and E2Bmin RNA.....	28
2.1.5 Crystallization of She2p in complex with <i>ASH1</i> E2Bmin-21 variants	29
2.1.6 Can She2p crystals be soaked with RNA?	30
2.1.7 Exploratory NMR experiments with She2p	31

TABLE OF CONTENTS

2.2	Functional analysis of mRNP assembly	32
2.2.1	She2p binds to zipcode elements with nanomolar affinity	32
2.2.2	She2p has considerable affinity to stem-loop containing RNAs	34
2.2.4	She2p has low affinity to a short zipcode-consensus sequence	35
2.2.5	Identification of conserved regions in She2p for mutational analysis	37
2.2.6	The basic helical hairpin motif provides a general binding platform for stem-loop containing RNA	39
2.2.7	The upper uncharged surface region is also required for recognition of stem-loop containing RNAs	39
2.2.8	Helix E is required for recognition of <i>ASH1</i> zipcodes	41
2.2.9	The She2p C-terminus also contributes to <i>ASH1</i> -mRNA binding	42
2.2.10	Reduction of basic RNA binding abolishes bud-tip localization of She2p <i>in vivo</i>	43
2.2.11	Selectively impaired <i>ASH1</i> -mRNA binding results in mislocalization of She2p <i>in vivo</i>	44
2.2.12	Impaired <i>ASH1</i> -mRNA binding by She2p results in mislocalization of <i>ASH1</i> mRNA <i>in vivo</i>	46
2.2.13	Impaired RNA binding by She2p affects translocation-complex assembly	46
2.2.14	She2p forms homo-oligomers through the upper surface region	48
2.2.14.1	Dynamic Light Scattering indicates large She2p particles	49
2.2.14.2	Analytical ultracentrifugation identifies She2p decamers that assemble through the upper surface region	50
2.2.14.3	Electron microscopy suggests She2p oligomers smaller than a decamer	53
2.2.15	She2p is tetrameric at physiological protein concentrations	53
2.2.16	Small Angle X-ray Scattering (SAXS) studies identify She2p tetramers	56
2.2.16.1	Identification of suitable buffer conditions for SAXS measurements	56
2.2.16.2	SAXS confirms tetramer formation in solution	58
2.2.17	SAXS indicates that She2p forms an elongated tetramer	61
2.2.18	Low resolution models of wild-type She2p	66
2.2.19	Attempt to obtain an atomic model of tetrameric She2p	68
3	DISCUSSION	69
3.1	She2p binds RNA through a composite binding site	69
3.2	She2p tetramers are the functional units	76
3.3	She2p-tetramer formation allows simultaneous binding of different RNAs	80
3.4	How are She2p-dependent priming complexes stabilized? - Implications for She2p-dependent mRNP assembly in <i>S. cerevisiae</i>	82
3.5	She2p binds RNA in multiple modes	85
4	MATERIALS AND METHODS	87
4.1	Consumables and chemicals	87
4.2	Oligonucleotides	87
4.2.1	DNA oligonucleotides	87
4.2.2	RNA oligonucleotides	88
4.2.3	DNA oligonucleotides to generate templates for <i>in vitro</i> transcription	89

TABLE OF CONTENTS

4.3	Plasmids	90
4.3.1	Plasmids for protein expression in <i>E. coli</i>	90
4.3.2	Yeast plasmids	91
4.4	Bacterial strains	91
4.5	Yeast strains	91
4.6	Media and supplements	92
4.7	Antibodies	93
4.8	Bioinformatic tools and software	93
4.9	Cloning, expression and purification methods	94
4.9.1	Cloning	94
4.9.2	Recombinant protein expression in <i>E. coli</i>	94
4.9.3	Isotopic ¹⁵ N-labeling for NMR spectroscopy	95
4.9.4	Purification of recombinant She2p variants	95
4.10	Methods for protein analysis	96
4.10.1	Western blotting	96
4.10.2	Dynamic Light Scattering	97
4.10.3	Analytical ultracentrifugation	97
4.10.4	Analytical size-exclusion chromatography	98
4.10.5	Monoclonal antibodies	98
4.11	Biochemical analysis of She2p:RNA interactions	99
4.11.1	<i>In vitro</i> transcription and RNA purification	99
4.11.2	Isotopic labeling of RNA oligonucleotides	100
4.11.2.1	5'-end labeling	100
4.11.2.2	3'-end labeling	100
4.11.3	Electrophoretic mobility shift assay (EMSA)	101
4.11.4	RNA filter-binding assay	101
4.11.5	Surface Plasmon Resonance	101
4.12	<i>In vivo</i> analysis of protein function	102
4.12.1	Transformation and cultivation of yeast cells	102
4.12.2	Preparation of yeast analytical whole-cell extracts	102
4.12.3	Co-immunoprecipitation	102
4.12.4	Fluorescence microscopy	104
4.12.4.1	Preparation of yeast spheroplasts	104
4.12.4.2	Indirect immunofluorescence	104
4.13	Methods for structural analysis	104
4.13.1	NMR data acquisition	104
4.13.2	Crystallization of She2p:RNA complexes	105
4.13.3	Small Angle X-ray Scattering (SAXS) experiments and data processing	105
5	REFERENCES	107

TABLE OF CONTENTS

6	APPENDIX.....	120
6.1	Secondary-structure predictions of zipcode elements and of unrelated stem-loop containing RNAs.....	120
6.2	RNA oligonucleotides used for co-crystallization experiments with She2p-(6-246)- Δ Cys	121
6.3	2D ^1H - ^{15}N -HSQC spectrum of wild-type She2p.....	121
6.4	She2p-sequence alignment.....	122
6.5	Assessment of <i>ASH1</i> -mRNA localization by <i>in situ</i> hybridization	123
6.6	Sedimentation profile of wild-type She2p as obtained in sedimentation-velocity experiments	123
	CURRICULUM VITAE.....	124
	PUBLICATIONS	125
	ACKNOWLEDGEMENTS	126

Summary

Selective translocation of *ASH1* mRNA during the budding of *Saccharomyces cerevisiae* represents the most comprehensively understood example of mRNA localization. The three core components of the *ASH1*-translocation complex - the RNA-binding protein She2p, the adapter protein She3p, and the myosin-motor protein Myo4p - assemble together into an mRNP and transport the RNA cargo to the daughter cell. Besides *ASH1* mRNA, more than 30 mRNAs are known to localize in a She2p-dependent manner. Each mRNA contains distinct zipcode elements that need to be specifically recognized by She2p. The present study dealt with the question, how the RNA-binding protein She2p selectively recognizes its target mRNAs and how RNA binding influences the assembly of functional mRNPs.

Biochemical studies revealed that She2p binds with nanomolar affinity to different zipcode elements from localizing mRNAs. Surprisingly, She2p showed significant binding to a subset of unrelated stem-loop containing RNAs, raising the question how She2p discriminates between target and non-target RNAs. Mutational analyses identified the protruding α -helix E and the very C-terminus of She2p to be required for specific *ASH1*-mRNA binding, but not for binding to unrelated stem-loop containing RNAs. In contrast, the basic unspecific recognition of RNA stem-loop structures was assigned to the basic helical hairpin RNA-binding motif. Selective impairment of specific RNA binding by mutation of the α -helix E or the very C-terminus results in loss of *ASH1*-mRNA and She2p localization *in vivo*. In addition, the formation of stable transport complexes is disrupted, demonstrating the requirement of correct target-mRNA recognition by She2p for the assembly of translocation-competent mRNPs. In the previously determined crystal structure of She2p, a dimer was suggested to be the functional oligomeric state in solution. The present study, however, revealed that She2p forms larger oligomers through its upper uncharged surface area. Two She2p dimers associate in a head-to-head orientation to form an elongated tetrameric complex. Along the opposing flat sides of this tetramer, two extended continuous RNA-binding surfaces are generated by the basic helical hairpin motif, the protruding helix E, and the very C-terminus of She2p. Such a composite RNA-binding motif on She2p tetramers is indispensable for RNA binding and specific mRNP assembly *in vivo*. With its combination of static and flexible structural elements She2p appears to be well designed to specifically recognize a number of different localizing mRNAs by applying multiple binding modes.

1 Introduction

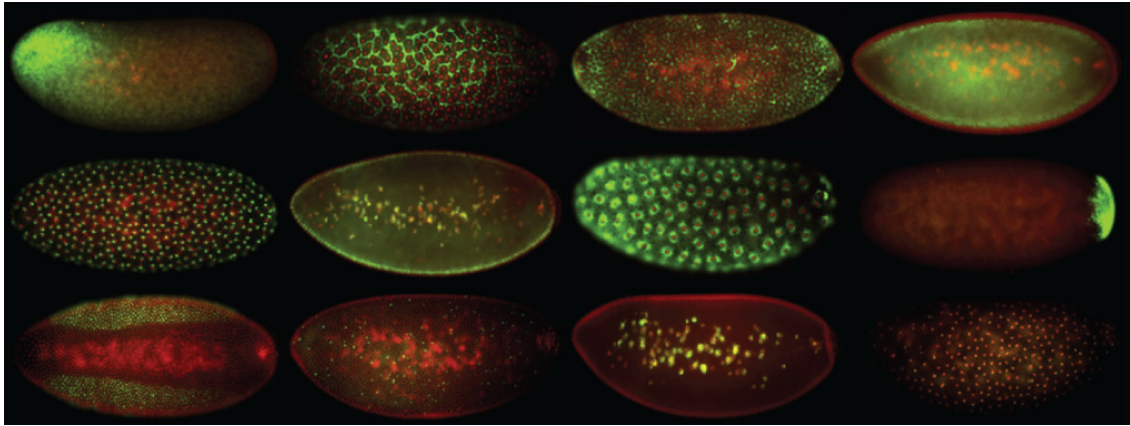


Figure 1: mRNA localization is a universal mechanism to control cellular architecture and function. Analysis of about 25 % of the mRNAs encoded by the *Drosophila* genome by high-resolution fluorescent *in situ* hybridization revealed that among those RNAs 71 % of expressed transcripts are subcellularly localized (Lecuyer et al., 2007). The illustration shows a selection of the observed localization patterns in early developmental stages of *Drosophila* embryogenesis. The anterior pole is to the left, the posterior pole to the right. mRNA is stained in green, nuclei are stained in red. Image was taken from Martin and Ephrussi, 2009.

A milestone publication by the laboratory of Henry M. Krause showed that the majority of transcripts expressed during early embryogenesis of *Drosophila* display specific subcellular localizations (Lecuyer et al., 2007). Besides the most prevalent sub-embryonic localization patterns, they identified a large number of mRNAs exhibiting novel distribution patterns. For example, they observed varieties in apico-basal, membrane-associated or nuclei-associated patterns. In Figure 1, a selection of observed localization patterns is shown. Together with genetic approaches (Tenenbaum et al., 2000; Ule et al., 2003), these systematic genome-wide analyses have enabled global surveys of the mRNA localization dynamics. Importantly, these studies also revealed tight correlations between mRNA localization and the distribution and function of the encoded proteins, suggesting an important role of transcript localization in organizing cellular architecture.

The following introduction begins with a summary of the mechanisms by which RNAs can become localized. Next, the assembly and translocation of an RNA transport particle will be explained, followed by a presentation of selected examples for localized mRNAs in somatic cells, in oocytes, and in embryos of higher eukaryotes. Finally, *Saccharomyces cerevisiae*, the best-characterized model organism for studying mRNA translocation, will be introduced and the principles of mRNA localization in yeast will be explained in detail.

1.1 mRNA localization – a common mechanism for targeting proteins

The discovery of the signal peptide by Günter Blobel and colleagues suggested that proteins are targeted to specific regions in a cell exclusively following translation (Blobel and Dobberstein, 1975). However, over the last 20 years it became clear that subcellular localization of messenger RNA also serves as a way of post-transcriptional gene regulation. Precise spatial and temporal control of the mRNA-localization process permits localized translation at discrete sites within cells. This specific positioning of cytoplasmic determinants results in cellular polarization and functional compartmentalization. It thereby contributes to many key-biological processes as body-axis specification, cell-fate determination, motility, and synaptic plasticity (Kloc et al., 2002; St Johnston, 2005). Why should cells regulate protein distribution by localizing its mRNA? An important reason is the possibility to spatially restrict protein synthesis within the cytoplasm. This allows cells to respond rapidly to local stimuli and, by this means, to regulate gene expression independently in different parts of the cell. Furthermore, mRNA localization not only targets the protein to the site of destination but also prevents its synthesis elsewhere. This is important if the protein might have toxic or deleterious effects on other cellular compartments. Additionally, a localized mRNA can be translated “in place” multiple times resulting in high local protein concentration. This process is more economic than transporting every single copy of the protein to a distinct site.

Asymmetric sorting of mRNAs can be achieved by several processes. The simplest, but very rare mechanism to localize an mRNA is by its local synthesis. This has been described for the mRNAs encoding the δ - and ϵ -subunits of the acetylcholine receptor that are exclusively transcribed in nuclei close to the neuromuscular junctions in mammalian myofibers (Brenner et al., 1990; Simon et al., 1992). The local transcription thus targets the mRNAs to the synapses, their later site of function. Secondly, mRNAs can be asymmetrically sorted by localized protection of degradation. This mechanism has been shown best for the *heat-shock protein-83* (*hsp83*) mRNA, which restricts *hsp83* mRNA to the posterior pole plasm in the egg of *Drosophila melanogaster* (Ding et al., 1993). Interestingly, this process requires two independent regions within the 3' untranslated region of the mRNA. One element targets the mRNA for destruction within all regions of the egg and a second element protects and stabilizes correctly localized mRNAs at the posterior pole (Bashirullah et al., 1999). Thirdly, mRNAs can become localized to specific subcellular regions by passively diffusing through the cytoplasm and subsequent entrapment by a localized anchor. This mechanism is known to

translocate *Drosophila nanos* mRNA late in oogenesis from the nurse cell to the posterior pole of the oocyte. At the posterior pole *nanos* localizes by association with the assembled pole plasm and by actin-dependent anchoring (Forrest and Gavis, 2003). This mechanism has also been proposed for *Xcat2* mRNA from *Xenopus laevis*. In germ cell development, *Xcat2* mRNA localizes to the developing germ plasm by random diffusion through the cytoplasm and successive immobilization through association with the endoplasmic reticulum in the mitochondrial cloud (Chang et al., 2004). Lastly, the most-prominent and probably best-characterized mechanism for asymmetric distribution of mRNA is the active and directed transport (St Johnston, 2005). This multistep process starts with the recognition of *cis*-acting elements within the mRNA, termed “localization elements” or “zipcodes”, by *trans*-acting factors. These are usually dedicated RNA-binding proteins, which then recruit specific motor-protein complexes. Dependent on the type of the motor protein, the mRNA-transport particle (mRNP) is transported along microtubules (dynein, kinesin) or actin filaments (myosin). At the site of destination, the mRNA is anchored and the encoded protein is translated. The process of directed mRNA transport will be explained in detail in the following chapter (1.2).

1.2 Assembly and translocation of an mRNA-transport particle

From the first discoveries of localized transcripts about 30 years ago until now, we know more and more details of the mechanisms underlying localization processes. However, despite the dramatically increased understanding, the initially proposed general principles how mRNAs are actively localized, have virtually not changed. In the early 1990s, an ordered pathway for unidirectional movement of “particles” has been proposed and is sketched in Figure 2 (Wilhelm and Vale, 1993). With respect to additional information we now have on each of the described steps, this scheme is still valid.

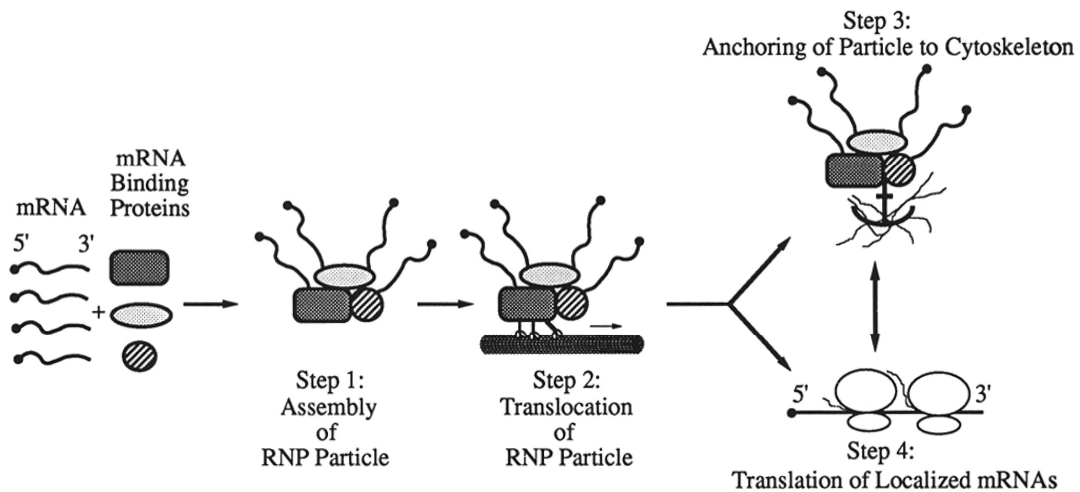


Figure 2: Pathway for mRNP translocation. In the nucleus, dedicated RNA-binding proteins recognize specific *cis*-acting elements in their target mRNA, which reside usually in the 3' UTR and often fold into stem-loop structures. During splicing, additional factors assemble on the nuclear mRNP (step 1), which is then exported to the cytoplasm. In the cytoplasm, the mRNP is remodeled. Additional proteins like translational repressors and proteins linking the mRNP to motor proteins join the complex (step 1). Dyneins and kinesins transport the mature mRNP along microtubules, whereas myosins translocate their cargo along actin microfilaments to the final site of destination (step 2). After anchoring the mRNP at cytoskeletal elements (step 3), the translational repression is released and the localized mRNA is spatially translated (step 4). The figure was taken from Wilhelm and Vale, 1993.

1.2.1 Assembly of the mRNP (Step1 in Figure 2)

The first step in mRNP assembly already occurs in the nucleus when a dedicated *trans*-acting factor specifically binds most likely co-transcriptionally to *cis*-acting elements within the mRNA target. It is suggested that the cytoplasmic fate of localizing mRNAs is predestined by factors acquired in the nucleus (Giorgi and Moore, 2007).

Cis-acting elements

In order to become specifically localized to subcellular regions, mRNAs contain so called *cis*-acting localization elements or “zipcodes”. These elements comprise the cellular “address” of the transcript and are recognized by specific *trans*-acting factors. Zipcodes are most frequently found in the 3' untranslated region (UTR) of the transcript, although in some cases they also reside in the 5' UTR or within the coding region (Jambhekar and DeRisi, 2007). In general, zipcodes function independently of nearby RNA sequences. Zipcodes fused to reporter RNAs direct a subcellular distribution of the reporter RNA similar to that observed for the native mRNA (Jambhekar and DeRisi, 2007). In some instances, zipcodes have been shown to function in different cell types, suggesting a certain redundancy when recognized by shared,

homologous RNA-binding proteins (Bullock and Ish-Horowicz, 2001). Zipcodes are very heterogeneous in their nature and complexity and thus, are very difficult to predict (Hamilton and Davis, 2007). The length of zipcodes can range from only five to six nucleotides (e.g. the VM1 and E2 elements in the 3' UTR of the *Xenopus Vg1* mRNA) up to 1 kb (e.g. the 3' UTR of *MBP* mRNA) (Ainger et al., 1997; Deshler et al., 1997). In case of *MBP* mRNA (see also below, chapter 1.3.2), it is only the primary sequence of the 11-nucleotides zipcode A2RE, which directs the localization of *MBP* mRNA to oligodendrocyte processes (Hoek et al., 1998). However, localization elements usually form distinct secondary structures, most commonly stem-loop structures that are critical for localization. In some mRNAs, like in *Xenopus Vg1* mRNA, multiples copies of the zipcode are clustered and act synergistically to localize the mRNA (Lewis et al., 2004). *Drosophila nanos* mRNA contains four zipcode elements in its 3' UTR, each of which localizes *nanos* to the posterior pole of the oocyte. However, all four elements together localize *nanos* with high efficiency to its final site (Gavis et al., 1996). Similarly, *ASH1* mRNA from *Saccharomyces cerevisiae* contains four distinct zipcodes, each of which are sufficient for localization but are altogether required for full levels of *ASH1* localization (for further details on mRNA localization in *S. cerevisiae*, see chapter 1.5) (Chartrand et al., 1999; Gonzalez et al., 1999). Such synergistic effects are believed to facilitate the association of *trans*-acting factors and the formation of a translocation-competent mRNP.

***Trans*-acting factors**

Trans-acting factors play an important role in directing and regulating mRNA localization, since they recognize the zipcodes within the mRNA, recruit motor proteins, and link the mRNA target to cytoskeletal tracks. However, the molecular mechanisms underlying the coupling of mRNA to motor complexes are only poorly understood. *Trans*-acting factors have been identified by genetic screens for genes involved in mRNA localization and by affinity purification of proteins associated with known localization elements (Martin and Ephrussi, 2009). One of the most-prominent *trans*-acting factors involved in mRNA localization is Staufen. Staufen was first identified in *Drosophila melanogaster* because of its role in body-axis formation, which reflects its function in *oskar* and *bicoid* mRNA localization (St Johnston et al., 1991) (see also chapter 1.4.2). When injected into early embryos, Staufen binds to stem-loop structures in the *bicoid* 3' UTR to form mRNPs that move in a microtubule-dependent manner (Ferrandon et al., 1994). *Drosophila* Staufen contains five conserved double-stranded RNA-binding domains (dsRBD), each of which can bind to dsRNA (Ferrandon et al., 1994; St Johnston et al., 1992). Interestingly, the Staufen dsRBDs are individually necessary for

different transport pathways. For instance, dsRBD5 mediates the actin-dependent localization of *prospero* mRNA, whereas dsRBD2 is necessary for microtubule-dependent *oskar* localization (St Johnston, 2005). Thus, different domains in Staufen connect the Staufen:mRNA complexes to both actin- and microtubule-dependent transport pathways (Micklem et al., 2000). However it is unclear, how the assembly of such diverse mRNPs is regulated. Another well known conserved *trans*-acting factor is the zipcode binding protein 1 (ZBP1). ZBP1 has been identified by affinity-purification experiments of proteins interacting with the β -*actin*-mRNA zipcode in chick fibroblasts (Ross et al., 1997). ZBP1 contains two RNA recognition motifs (RRM) and four hnRNP K homology (KH) domains, for which distinct functions have been identified (Farina et al., 2003). The RRM s are required for the localization of the β -*actin* mRNP itself, whereas two out of four KH domains specifically bind to the β -*actin* zipcode and connect the assembled mRNP to the actin cytoskeleton (Farina et al., 2003). In addition, ZBP1 is implicated in the microtubule-dependent transport of β -*actin* mRNA in neurons (Tiruchinapalli et al., 2003; Zhang et al., 2001). In analogy to Staufen, ZBP1 is able to couple its target mRNA to different types of cytoskeletal elements, but additional regulatory factors for β -*actin* mRNP assembly are largely unknown. ZBP1 homologues in *Xenopus laevis* (Vg1RBP/Vera, see chapter 1.4.1) and in mammalian cells (IMP1) have both been shown to function in mRNA localization. This example nicely demonstrates the conservation of *trans*-acting factors among cell types and species (Atlas et al., 2004; Havin et al., 1998). However, another well-characterized *trans*-acting factor, the RNA-binding protein She2p from *S. cerevisiae*, lacks any homologous counterparts in higher eukaryotes (Niessing et al., 2004). She2p is responsible for the transport of at least 24 mRNAs (Oeffinger et al., 2007; Shepard et al., 2003; Takizawa and Vale, 2000) suggesting that not only highly conserved *trans*-acting factors with defined RNA-binding motifs are able to selectively recognize their targets. Further details on She2p-dependent mRNA localization are described in chapter 1.5.

1.2.2 Motor-driven translocation of a translationally silenced mRNP along cytoskeletal tracks (Step 2 in Figure 2)

Following export to the cytoplasm, the mRNP usually undergoes remodeling. Cytoplasmic factors join the RNA:protein complex, whereas nuclear factors might dissociate (Martin and Ephrussi, 2009). Assembled mRNPs associate with motor proteins and travel along microtubules or actin microfilaments, the railways for translocating mRNAs in the cytoplasm. Three main families of motor proteins exist: the dynein, kinesin, and the myosin families

INTRODUCTION

(Vale, 2003). Most mRNA translocation processes involve dynein and kinesin motors, which move along polarized microtubules. For instance, *Drosophila bicoid* mRNA is transported in a dynein-dependent manner to the anterior pole of the oocyte, whereas *oskar* mRNA localizes to the posterior pole requiring the kinesin-1 motor protein (Duncan and Warrior, 2002; Januschke et al., 2002; Zimyanin et al., 2008). Similarly, mRNAs translocate along microtubules in *Xenopus* oocytes coordinated by multiple kinesin motors (Messitt et al., 2008). In comparison, less mRNA-localization processes are known that require myosin-motor proteins and polarized actin cables. Examples for actin-dependent transport processes are the localization of β -actin mRNA in migrating fibroblasts as well as mRNA localization by the She machinery in *S. cerevisiae* (see also chapters 1.3.1 and 1.5). In yeast, the core components of the translocation complex are known and the interactions between the RNA-binding protein She2p, the linker protein She3p, and the motor protein Myo4p are well defined (see chapter 1.5) (Böhl et al., 2000; Jansen et al., 1996; Long et al., 2000; Takizawa and Vale, 2000). However, in other systems only the involvement of different motor proteins has been clearly established, while the connections between the zipcode-interacting proteins and motors are poorly understood (Bullock, 2007).

Especially in neuronal cells it was shown that mRNPs assemble into heterogeneous granules. Besides several mRNAs, these granules also contain factors of the translation machinery including ribosomal subunits and factors involved in translational regulation (see also chapter 1.2.3) (Kiebler and Bassell, 2006). It is generally assumed that localizing mRNAs are translationally silenced during the transport through the cytoplasm (Besse and Ephrussi, 2008). This is supported by the fact that proteins encoded by the localizing mRNAs accumulate exclusively at the sites of destination. Translational repressors were found to associate with localizing mRNAs (Besse and Ephrussi, 2008). Most frequently, translation initiation, the rate-limiting step in translation, is regulated by targeting the eukaryotic translation initiation complex eIF4F. For instance, this has been shown for the *Drosophila* Cup protein, which simultaneously binds to eIF4E and *oskar*-associated Bruno protein thus preventing premature translation of *oskar* mRNA (Nakamura et al., 2004). Similarly, in *S. cerevisiae* *ASH1*-mRNA translation is repressed by association of the RNA-binding proteins Khd1p and Puf6p (see chapters 1.2.3 and 1.5) (Deng et al., 2008; Paquin et al., 2007).

1.2.3 Anchoring of the mRNP and translation of localized mRNAs (Steps 3 and 4 in Figure 2)

At the place of destination, the mRNP needs to be anchored and translation must be activated. The actin cytoskeleton was shown to be important for anchoring of a number of localizing mRNAs, including *ASH1* mRNA in yeast, β -actin mRNA in fibroblasts, *arc* mRNA in neurons, *bicoid* mRNA in *Drosophila* oocytes and embryos, and *Vg1* mRNA in *Xenopus* (Beach et al., 1999; Farina et al., 2003; Huang et al., 2007; Takizawa et al., 1997; Weil et al., 2008; Yisraeli et al., 1990). After anchoring, the translational repression must be abrogated. This is commonly achieved by reducing the affinity of repressor complexes to their target mRNAs, for instance by spatially restricted phosphorylation of RNA-associated proteins (Besse and Ephrussi, 2008). Kinase-based regulation of RNA-binding proteins most likely represents a common mechanism for translational regulation, although it was observed for only a few proteins. In chick fibroblasts, phosphorylation of ZBP1 by Src kinase decreases ZBP1's affinity to the bound β -actin zipcode and results in spatially restricted synthesis of β -actin protein (Hüttelmaier et al., 2005). In yeast, the two translational repressors of *ASH1* mRNA, Puf6p and Khd1p, are phosphorylated by casein kinase II (CK2) and the type I casein kinase Yck1p, respectively. This leads to dissociation of Puf6p and Khd1p from *ASH1* mRNA and activation of localized translation (Deng et al., 2008; Paquin et al., 2007) (see also chapter 1.5).

1.3 mRNA localization in somatic cells

As mentioned above, localization of mRNA is a global mechanism to restrict protein synthesis to specific subcellular regions. This and the following chapter 1.4 aim to present selected, well-understood examples of localized mRNAs in higher eukaryotes. In polarized somatic cells such as fibroblasts or nerve cells, localization of mRNA often contributes to cell motility and the development of cellular morphology.

1.3.1 mRNA localization in fibroblasts

In migrating chick and mammalian fibroblasts, β -actin mRNA localizes to the leading edge of lamellipodia where increased levels of the β -actin protein are required for cytoskeleton-mediated movement (Figure 3 a) (Condeelis and Singer, 2005; Lawrence and Singer, 1986). β -

actin mRNA localization in chick fibroblasts is initiated by co-transcriptional association of the zipcode binding protein 2 (ZBP2), a homologue of the human splicing factor KSRP, with the nascent transcript (Gu et al., 2002; Pan et al., 2007). This interaction then facilitates subsequent binding of the shuttling protein ZBP1 to the 54-nucleotide zipcode in the 3' UTR of *β-actin* mRNA (Kislauskis et al., 1994; Pan et al., 2007; Ross et al., 1997). ZBP1 stays associated with the mRNA during the localization process and prevents premature translation by blocking translation initiation. At the final destination translational repression is released by phosphorylation of ZBP1 (see also chapter 1.2.3) (Hüttelmaier et al., 2005). The cooperative action of ZBP1 and ZBP2 nicely shows that the cytoplasmic localization of an mRNA is already triggered by the association of nuclear factors, thereby regulating the formation of a functional mRNP (Martin and Ephrussi, 2009). Additionally, these sequential events reflect the temporal and spatial control of *β-actin* mRNA localization that results in exclusive accumulation of the *β-actin* protein at lamellipodia.

1.3.2 mRNA localization in the neuronal system

Neurons are highly polarized and compartmentalized cells. Formation and maintenance of a functioning nervous system relies in part on the localization of mRNAs to defined regions in the neuron, including axons and dendrites. This eventually allows local translation in the extremities of the neuronal cells (Job and Eberwine, 2001). Purification of RNA granules from mouse brain that associate with the kinesin-motor protein KIF5 resulted in the identification of *CamKIIa* and *arc* mRNAs and about 40 proteins (Kanai et al., 2004). Among them were proteins previously known to function in mRNA localization in neurons, as Staufen1, the Fragile X Mental Retardation Protein (FMRP), and Pur- α as well as components of the translation machinery. However, also new proteins were found to be involved in dendritic RNA transport, including hnRNP U and polypyrimidine tract binding protein-associated splicing factor (PSF). This study suggests that many components of RNA granules may not be essential for the localization process itself, but may rather be structural components involved in mRNP assembly and stabilization. The *CamKII* gene encodes the calcium-dependent calmodulin protein kinase II, a serine-threonine kinase highly expressed in adult neurons composed of α - and β -subunits. *CamKIIa* mRNA encodes the α -subunit of CamKII and localizes to dendrites dependent on at least three localization elements residing in the 3' UTR of the transcript (Dahm et al., 2007). However, the *trans*-acting factors binding to the *CamKIIa* mRNA are still largely unknown. In mammalian oligodendrocytes, *myelin basic protein* (MBP)

mRNA localizes to myelinating compartments that enclose neuronal axons (Figure 3 b). Interestingly, directed transport of *MBP* mRNA depends on two distinct localization elements (Ainger et al., 1997; Munro et al., 1999). One of these is called RNA trafficking signal (RTE) or A2RE, is only 11 nucleotides long and is bound by the heterogeneous ribonucleoprotein A2 (hnRNP A2). The second localization element, the RNA localization region (RLR), spans much of the 3' UTR and is predicted to form extensive secondary structures with multiple stem loops. The first zipcode directs the transport of the *MBP* mRNA out of the soma whereas the latter targets the mRNA into the myelinating compartment.

1.4 mRNA localization in oocytes and developing embryos

Body patterning during the development of multicellular organisms essentially relies on the localization of mRNA to subcellular regions. This chapter aims to show some of the best-understood examples of localized mRNAs in *Xenopus* and *Drosophila* oocytes and embryos.

1.4.1 mRNA localization in *Xenopus laevis* oocytes

In *Xenopus* oocytes, mRNAs are asymmetrically sorted to the animal or vegetal pole and thereby contribute to the definition of the germ cell lineage and the development of the animal/vegetal axis. Two indispensable pathways exist, through which mRNA localization is temporally restricted to early (stages I and II) and late (stages III-V) developmental stages (Kloc and Etkin, 1995). In the early pathway or METRO (messenger transport organizer), germ granules, mRNAs involved in germ cell specification, and germ-line specific mitochondria are transported to a specialized structure, which is known as the Balbiani body or the mitochondrial cloud. It serves as a vehicle to localize RNAs at the vegetal pole (Kloc and Etkin, 1995). Mitochondrial-cloud localization elements (MCLEs) that direct the RNA into the mitochondrial cloud have been discovered in the 3' UTR of *Xcat2* mRNA and within the non-coding *Xlirts* RNA (Allen et al., 2003; Kloc et al., 1993; Zhou and King, 1996). Additionally, *Xcat2* mRNA contains the germinal granule localization element, which determines *Xcat2* mRNA localization to the germ-plasm region within the mitochondrial cloud (Figure 3 c) (Kloc et al., 2000). The localization of RNAs to the mitochondrial cloud is believed to be independent from cytoskeletal elements and occurs most likely through a diffusion/entrapment mechanism (see chapter 1.1) (Chang et al., 2004; Kloc et al., 1996). However, there is recent evidence that the ATP availability in the oocyte regulates *Xcat2*-

mRNA localization and that the motor protein kinesin II is involved in this process (Heinrich and Deshler, 2009). The late pathway during mid and late oogenesis localizes mRNAs such as *Vg1* and *FatVg* to the vegetal cortex (Kloc and Etkin, 2005). This process depends on microtubules and is believed to involve kinesin-motor proteins. Additionally, a unique endoplasmic-reticulum (ER) subdomain forms between the oocyte nucleus and the vegetal pole in which localizing mRNAs accumulate. It is still not clear if the ER serves as a matrix for the formation of microtubule tracks or if localizing RNAs bind to ER vesicles and are co-transported along microtubules to the vegetal pole (Deshler et al., 1997; Kloc and Etkin, 1998). Multiple redundant localization elements within the 3' UTRs of *Vg1* mRNA (see also chapter 1.2.1) and several *trans*-acting factors associated with them have been identified (Kloc and Etkin, 2005). Among these factors is the protein Vg1RBP/Vera that binds to zipcodes in the *Vg1* 3' UTR through its KH domains and mediates association with microtubules (Deshler et al., 1998; Havin et al., 1998). Vg1RBP/Vera is the homologue of mammalian ZBP1, whereas another *Vg1* mRNA-binding protein, Vg1RBP71, was shown to be the homologue of ZBP2. This demonstrates a functional relevance of conserved *trans*-acting factors (Martin and Ephrussi, 2009).

1.4.2 Localization of *gurken*, *bicoid* and *oskar* mRNA during *Drosophila* oogenesis

Drosophila oogenesis represents a great system to investigate differential sorting of mRNAs within the same cell. It results in the selectively localized translation of proteins establishing the prospective body axis. At early stages in oogenesis, maternal mRNAs are transported from the nurse cells into the oocyte. For this transport process, a complex consisting of the microtubule-dependent motor dynein, and the proteins Egalitarian and Bicaudal-D is required (Bullock and Ish-Horowicz, 2001). Later, three of these early localized mRNAs, *gurken*, *bicoid*, and *oskar*, relocate within the oocyte in a dynein- (*gurken* and *bicoid* mRNA) and kinesin- (*oskar* mRNA) dependent fashion. *Gurken* encodes for the *Drosophila* homologue of transforming growth factor α (TGF α), which is required for sequential establishment of both the antero-posterior and the dorso-ventral axes (Gonzalez-Reyes et al., 1995; Neuman-Silberberg and Schupbach, 1993). *bicoid* and *oskar* mRNA localize to the anterior and posterior of the oocyte, respectively, thus patterning the antero-posterior axis (Berleth et al., 1988; Ephrussi et al., 1991; Kim-Ha et al., 1991). In addition to the *oskar* 3' UTR, several factors are required for *oskar* localization to the posterior pole. The splicing process of *oskar* mRNA

triggers the complex formation with components of the exon-junction complex, which accompany the mRNA to its final site (Hachet and Ephrussi, 2004). The cytoplasmic proteins Barentz, Hrp48 and Staufen also co-localize with *oskar* mRNA at the posterior pole and are supposed to be part of the localization complex (St Johnston, 2005). Localization of *bicoid* mRNA to the anterior pole of the oocyte occurs during different stages in oogenesis and involves several *cis*- and *trans*-acting factors (Figure 3 d). At late stages, *bicoid* localization to the anterior pole relies on the interaction of Staufen with stem-loop structures in the *bicoid* 3' UTR (Ferrandon et al., 1994; St Johnston et al., 1991). The encoded protein Bicoid is a transcription factor that forms a morphogenic gradient by diffusing from the anterior pole throughout the embryo (Ephrussi and St Johnston, 2004).

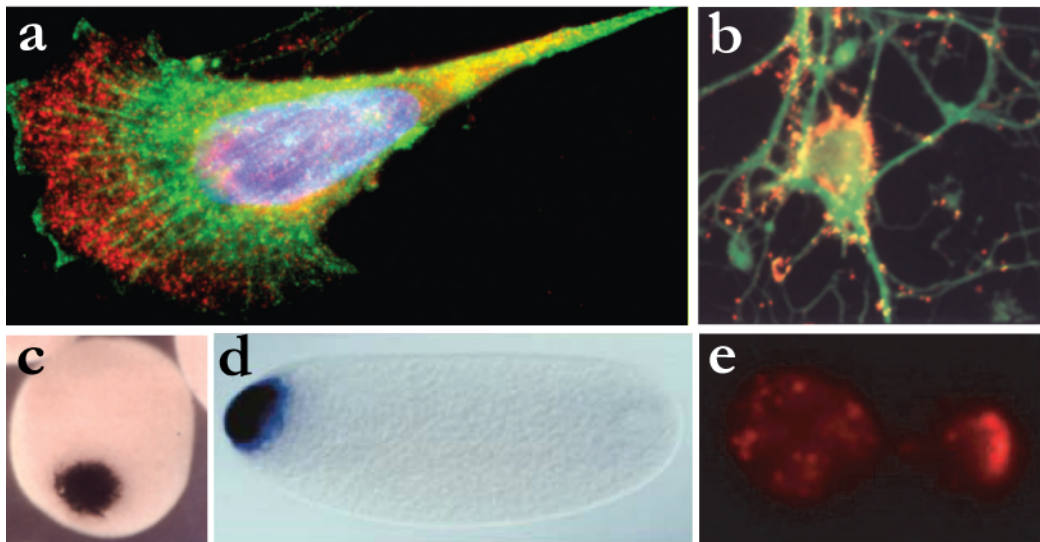


Figure 3: RNA localization is observed in various cell types and organisms. a) *β-actin* mRNA (red) localizes to the leading edge of a migrating fibroblast cell. *β-actin* protein is stained in green and the nucleus is stained in blue. **b)** *MBP* mRNA-transport particles (red) associate with microtubules (green) and localize to myelinating processes in oligodendrocytes. **c)** Localization of *Xenopus laevis* *Xcat2* mRNA to the vegetal cortex in a stage IV oocyte. **d)** In the *Drosophila* egg, *bicoid* mRNA is tightly localized to the anterior pole. **e)** In *Saccharomyces cerevisiae*, *ASH1* mRNA localizes to the bud tip in late anaphase of a dividing cell. Pictures were taken from Dahm and Kiebler, 2005; Ephrussi and St Johnston, 2004; St Johnston, 1995, 2005.

1.5 mRNA localization in *Saccharomyces cerevisiae*

As shown in previous chapters, in higher eukaryotes a significant number of mRNP components has been identified, but a precise function was assigned to only a few of them, reflecting the high complexity of these transport complexes (Martin and Ephrussi, 2009). In contrast, in the budding yeast *Saccharomyces cerevisiae*, the essential components of the *ASH1*

mRNA-translocation complex are known and have been genetically and biochemically characterized (Jansen et al., 1996; Martin and Ephrussi, 2009; Paquin and Chartrand, 2008). The initial genetic screen identified 5 genes (*SHE1-5*), whose gene products are involved in *ASH1*-mRNA translocation (Jansen et al., 1996). Three of them, the proteins She1p, She2p and She3p, were shown to represent the core components of the *ASH1* translocation complex, also referred to as the locosome (Böhl et al., 2000; Gonsalvez et al., 2004; Long et al., 2000; Takizawa and Vale, 2000). She1p/Myo4p is a type V myosin-motor protein, which transports the mRNP along the polarized actin network. The motor protein is connected via the adapter protein She3p to the RNA-binding protein She2p. She2p itself binds to the mRNA target and thereby links the cargo to the motor of the transport complex. She4p, a member of the UCS class of proteins, associates with Myo4p and is supposed to function in proper folding of the Myo4p-motor domain (Wesche et al., 2003). Lastly, Bni1p/She5p is a formin that is involved in actin filament organization, thus ensuring continuous transport of the *ASH1* mRNP (Beach et al., 1999; Evangelista et al., 2003). To date, the *ASH1* mRNP represents the best-characterized example for directed transport of a cell fate determinant to subcellular regions (Figure 3 e).

1.5.1 “Mating type switching” – the biological function of *ASH1*-mRNA localization

The life cycle of *S. cerevisiae* is characterized by diploid and haploid growth states. Under conditions of nutrient deprivation, *S. cerevisiae* undergoes meiotic cell division, which results in two haploid cells - one larger mother cell and a smaller daughter cell - of the opposite mating type “a” or “α”. To re-establish the diploid growth state, two cells of the opposite mating type mate to form an a/α diploid cell. Interconversion between “a” and “α” mating types in the haploid cell is achieved by the mother cell-specific asymmetric expression of the *HO* endonuclease. *HO* endonuclease initiates genomic rearrangements of the *MAT*-gene locus only in the haploid mother cell (Cosma, 2004). This phenomenon has been termed “mating type switching”. In the daughter cell, expression of *HO* endonuclease is repressed by the transcription factor Ash1p, which specifically binds to the *HO* endonuclease promoter and, thus, inhibits the transcription of the inducer of mating type switching (Bobola et al., 1996; Jansen et al., 1996; Sil and Herskowitz, 1996). Restriction of Ash1p to the daughter cell is achieved by specific translocation of the *ASH1* mRNA to the bud tip and its exclusive translation in the daughter cell.

1.5.2 *ASH1* and at least 23 additional bud-localized mRNAs

DNA-microarray analyses identified a total of 24 mRNAs, including *ASH1* mRNA, to specifically associate with tagged She2p-She3p-Myo4p complexes (Table 1) (Shepard et al., 2003; Takizawa and Vale, 2000). However, a recent study found that 41 mRNAs co-purify with Protein A-tagged She2p from native mRNPs suggesting that the number of She2p-dependent localized mRNAs in *S. cerevisiae* might be even larger (Oeffinger et al., 2007). It is interesting to note, that the majority of localized transcripts in budding yeast encodes for membrane or membrane-associated proteins.

Gene	Cell cycle regulation	Localization of encoded protein	Predicted protein function
<i>ASH1</i>	M	Bud nucleus	Transcription
<i>BRO1</i>	None	Punctae on vacuole	Stress transduction
<i>CLB2</i>	M	Nuclei, spindle poles	Cyclin B
<i>CPS1</i>	None	Cytoplasmic punctae	Carboxypeptidase
<i>DNM1</i>	S	Mitochondrial periphery	Mitochondrial fission
<i>EGT2</i>	M	Membranes, large-bud enriched	Cellulase
<i>ERG2</i>	M	Endoplasmic reticulum	Sterol isomerase
<i>IST2</i>	None	Bud plasma membrane	Transporter
<i>MID2</i>	None	Cell periphery, mother-bud junction	Membrane receptor
<i>MMR1</i>	M	Bud sites and tips, mother-bud junction	Unknown
<i>SRL1</i>	G1	Periphery of small buds	Unknown
<i>TPO1</i>	M	Bud plasma membrane	Polyamine transport
<i>WSC2</i>	S	Membranes, bud-enriched	Membrane receptor
<i>TAM41</i>	None	Mitochondria	Unknown
<i>IRC8</i>	M	Membranes, bud-enriched	Unknown
<i>YLR434C</i>	None	Mitochondria	Unknown
<i>TCB3</i>	G2	Membranes, bud-enriched	Lipid binding
<i>EAR1</i>	None	Endoplasmic reticulum	Unknown
<i>TCB2</i>	None	Membranes, bud-enriched	Unknown
<i>KSS1</i>	None	not defined	Mitogen-activated protein kinase
<i>LCB1</i>	None	Endoplasmic reticulum	Lipid synthesis
<i>MET4</i>	None	Nuclei	Transcription
<i>MTL1</i>	None	not defined	MID2-like
<i>YPL066W</i>	None	not defined	Unknown

Table 1: Overview of localized mRNAs in *S. cerevisiae*. The table shows yeast genes whose mRNAs are localized in a She protein-dependent manner and lists localization and function of the respective encoded proteins. Information given in this table was obtained from Shepard et al., 2003 and from the Saccharomyces Genome Database (<http://www.yeastgenome.org>).

INTRODUCTION

The best-characterized example of mRNA sorting in yeast is the localization of *ASH1* (asymmetric synthesis of *HO*) mRNA to the bud tip during late anaphase (Figure 3 e) (Paquin and Chartrand, 2008). In order to get recognized by the *trans*-acting factor She2p, *ASH1* contains four *cis*-acting elements, termed E1, E2A, E2B, and E3. Three of them (E1, E2A, E2B) are located in the coding sequence of *ASH1*, whereas the E3 zipcode comprises the termination codon and is mainly part of the 3' UTR (Chartrand et al., 2002; Chartrand et al., 1999; Gonzalez et al., 1999; Long et al., 1997; Takizawa et al., 1997). Interestingly, the *ASH1* zipcodes function redundantly, since each of them can individually localize a heterologous reporter RNA to the bud tip (Chartrand et al., 2002). However, the presence of all four zipcodes enhances the efficiency of the translocation process. Thus it provides another example for synergistic clustering of multiple elements (see chapter 1.2.1). As observed also in other systems, *ASH1* zipcodes show no obvious sequence similarity to each other, but are rather predicted to fold into distinct stem-loop structures that are recognized by She2p (Böhl et al., 2000; Long et al., 2000; Olivier et al., 2005). Huge effort was made to determine features in *ASH1* that are required for proper sorting. In two recent studies, nucleotides essential for She2p recognition have been identified by *in vivo* selection from libraries of partially randomized *ASH1* zipcodes and by non-homologous random recombination of zipcode-containing fragments of pools from known localized mRNAs, respectively (Jambhekar et al., 2005; Olivier et al., 2005). The resulting recognition motif consists of a loop-stem-loop RNA structure containing a highly conserved CGA triplet with a downstream stretch of adenosines in one loop and an oppositely located single conserved cytosine. Mutation of this motif resulted in *ASH1* mislocalization *in vivo*. Additionally, the distance between these nucleotides as well as their spatial orientation seems to be crucial for *in vivo* *ASH1*-mRNA localization. These studies suggested that She2p recognizes *ASH1* zipcodes on the basis of both primary sequence and secondary structure. Using their techniques, both laboratories identified so far unknown localization elements containing the She2p-recognition motif in other bud-localized mRNAs (Jambhekar et al., 2005; Olivier et al., 2005). Among them, they found single zipcodes in *IST2* and *EAR1*, and two zipcodes in *WSC2*, which were all shown to function in mRNA localization *in vivo*. Unfortunately, the major fraction of zipcodes within bud-localized mRNAs remains unknown.

1.5.3 She2p – the cargo-binding protein

Structural characterization of a significant number of RNA-binding proteins led to the identification of a variety of RNA-binding domains. Examples are the hnRNP K homology (KH) domain, the RNA-recognition motif (RRM), the double-stranded RNA-binding domain (dsRBD), or the glycine-rich domain (Dreyfuss et al., 2002). Surprisingly, the motif by which She2p binds its mRNA cargo could not be attributed to any of these RNA-binding motifs, suggesting that She2p represents a unique class of RNA-binding proteins with an unknown mechanism of RNA binding. The crystal structure of She2p revealed an almost exclusively α -helical protein, which has no similarity to previously described nucleic acid-binding proteins (Figure 4) (Niessing et al., 2004). This finding is not surprising, since She2p homologues are found only in related yeast species, thus emphasizing the novel protein fold. She2p folds into a single domain consisting of a bundle of five anti-parallel α -helices and an additional helix protruding at right angles from the middle of the bundle. In the crystal lattice, She2p forms a stable homodimer of 56 kDa in size. Mutation of residues Cys68 and Ser120 in the dimer interface disrupts dimer formation and results in inefficient binding to the *ASH1*-E3 zipcode *in vitro* and mislocalization of *ASH1* mRNA *in vivo*. Thus, She2p dimers were suggested to be the functional unit *in vivo* (Niessing et al., 2004).

In order to map the RNA-binding region in She2p, the electrostatic potential of the She2p surface was analyzed (Figure 4 d-f). The surface mainly consists of negatively charged (especially in the bottom region) and hydrophobic residues (in the upper surface region) with exception of a distinct positively charged region at the side of the protein. From previous studies it was known that several basic residues within the first 70 amino acids of She2p are important for *ASH1* mRNA binding (Gonsalvez et al., 2003; Kruse et al., 2002). Consistently, these residues match very well with the positively charged patch of She2p and lie in the putative RNA-binding region. Further *in vitro* RNA-binding studies and *in vivo* *ASH1* mRNA-localization experiments with She2p selectively mutated in this positively charged region confirmed the previous findings and resulted in the determination of a new RNA-binding motif. This motif, termed “basic helical hairpin”, consists of two antiparallel α -helices that are separated by a loop, and is located on each monomer of the She2p homodimer (Figure 4 a-c, labeled in yellow). It is thought that this local tertiary structure feature provides the primary anchor for *ASH1* mRNA binding (Niessing et al., 2004). Thus far, She2p’s affinity to zipcode-containing mRNA has been poorly analyzed. Only for one single localization element, the *ASH1* E3 zipcode, the equilibrium dissociation constant (K_D) was determined to be

INTRODUCTION

210 nM \pm 40 nM, indicating tight binding by She2p (Niessing et al., 2004). Since one She2p dimer contains two RNA-binding motifs, it is suggested that more than one RNA molecule can be bound by She2p. In fact, both a 1:1 and 1:2 ratio for the She2p dimer:RNA complex was found depending on the analyzed zipcode and may vary with the length of the zipcode (Niessing et al., 2004; Olivier et al., 2005). However, the affinity to remaining zipcodes as well as the She2p-binding mode remains elusive.

In order to bind its mRNA target, She2p is actively imported into the nucleus by interaction with the importin α Srp1p. Nuclear import of She2p is driven by a recently identified non-classical nuclear localization signal (NLS), which is located in a less conserved region at the side of each monomer (Figure 4 a-c, labeled in brown) (Shen et al., 2009). In addition to the basic helical hairpin, the upper surface region of She2p is remarkably interesting. This region is largely uncharged and contains a high degree of conserved, hydrophobic amino acids (Figure 4 c,e). Furthermore, this hydrophobic patch spans the surface of both monomers in the dimer (Niessing et al., 2004). Mutation of amino acids threonine 47 and leucine 130 into tyrosine significantly reduces She2p's affinity to the *ASH1*-E3 zipcode (Niessing et al., 2004). This observation is consistent with the isolation of a yeast strain bearing a leucine 130 to serine mutation that is defective for *ASH1*-mRNA localization (Gonsalvez et al., 2003). Because both amino acids are located at the upper surface region of She2p, this region might be involved RNA binding. However, it is also speculated that the upper surface serves as the She3p-interacting region (see below, chapter 1.5.4).

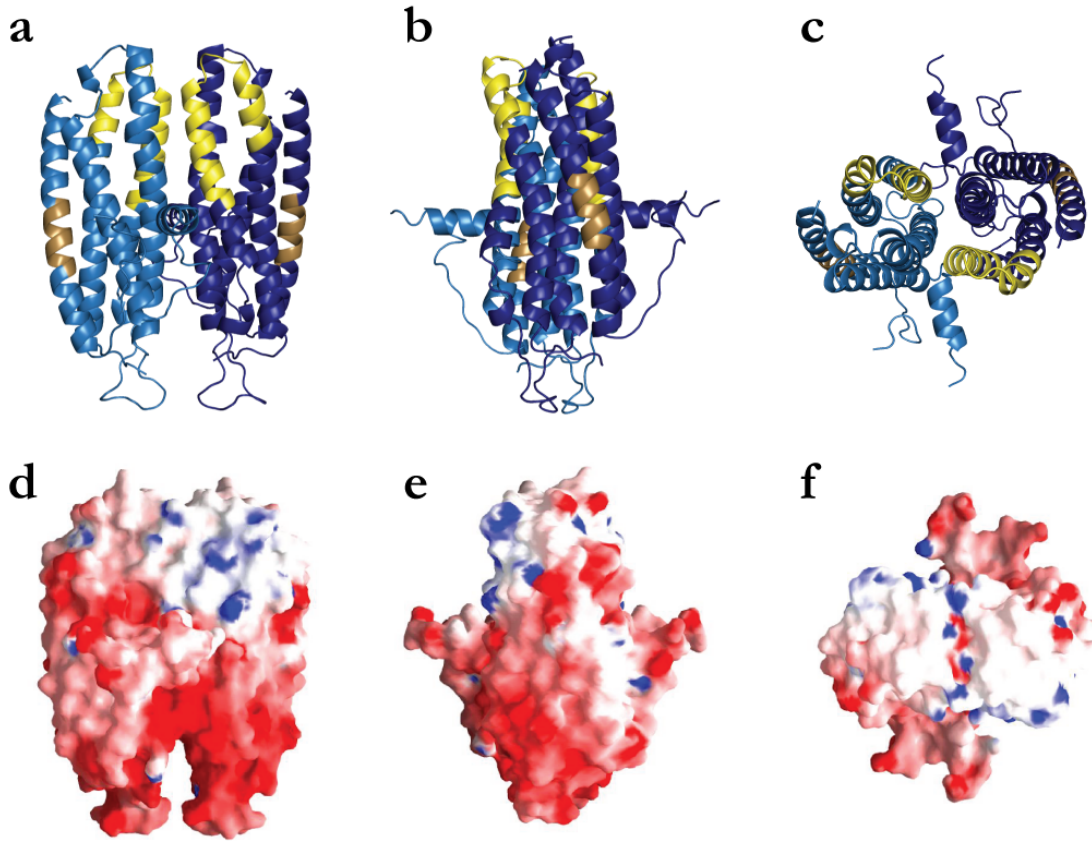


Figure 4: Crystal structure of the RNA-binding protein She2p. a)-c) Crystal structure of the She2p dimer in front view (a), side view (b) and top view (c). Each monomer is shown in light and dark blue, respectively. The basic helical hairpin RNA-binding motif is colored in yellow, the region comprising the nuclear localization signal is marked in brown. Relative to a), the structure in b) is rotated by 90 ° around the vertical axis, whereas c) is rotated by 90 ° around the horizontal axis with the top of the protein oriented to the reader. Images a)-c) were generated using PyMOL, DeLano Scientific, CA. d)-f) represent the electrostatic surface potential of the respective She2p cartoons shown in a)-c). The electrostatic potential is color-coded from red (negatively charged) to white (uncharged) to blue (positively charged) and was calculated using GRASP. Figures d)-f) are taken from Niessing et al., 2004.

1.5.4 She3p links the She2p:mRNA complex to the myosin-motor protein Myo4p

She3p serves as an adapter protein that connects the mRNA cargo to the actin cytoskeleton via its simultaneous interaction with the She2:mRNA complex and the unconventional type V myosin-motor protein Myo4p (Böhl et al., 2000; Long et al., 2000; Münchow et al., 1999; Takizawa and Vale, 2000). For the interaction with She2p, the C-terminal domain of She3p is responsible (Böhl et al., 2000; Long et al., 2000). Only in the presence of mRNA cargo, She3p interacts efficiently with She2p. Furthermore, She3p stabilizes the She2p:mRNA interaction (Böhl et al., 2000). These findings and the function of the upper hydrophobic surface region of She2p in mRNA binding and translocation (see chapter 1.5.3) led to the assumption that

She3p interacts with the hydrophobic surface of She2p in a mRNA-dependent manner, thus acting as a cargo sensor for mRNP assembly. The N-terminal domain of She3p binds to the coiled-coil region and the C-terminal tail in Myo4p, thereby forming a stable cytoplasmic complex (Böhl et al., 2000; Heuck et al., 2007; Long et al., 2000; Takizawa and Vale, 2000). However, in cells lacking She2p and thus, in the absence of correctly bound cargo, the She3p:Myo4p complex is equally distributed in mother and daughter cells (Gonsalvez et al., 2003; Kruse et al., 2002). Surprisingly, asymmetric sorting can be restored by artificially tethering a heterologous mRNA to She3p, which then localizes correctly to the bud (Long et al., 2000). This finding confirms that Myo4p is recruited to the mRNA target via She3p, which itself associates with She2p. Upon mRNP assembly, the complex is transported by Myo4p along actin microfilaments to the tip of the daughter cell, where it becomes anchored in an actin-dependent manner (Beach et al., 1999; Takizawa et al., 1997). It is important to note that the localizing mRNA needs to be translationally repressed during transport ensuring spatially restricted protein synthesis (see chapter 1.5.5).

In addition to mRNA transport, She3p and Myo4p function in the inheritance of cortical ER (Estrada et al., 2003; Schmid et al., 2006). This process is independent of She2p, but occurs in parallel with mRNA transport. ER inheritance can be blocked by deletion of the motor *MYO4* or by latrunculin-A treatment of yeast cells, which inhibits actin polymerization. This shows that ER transport directly depends on Myo4p (Estrada et al., 2003). As mentioned above in chapter 1.5.2, the major fraction of localized mRNAs in yeast encodes for membrane-associated proteins, which are synthesized at the rough ER. This suggests that the processes of mRNA localization and ER transport might be coordinated. Indeed, a study by the laboratory of Prof. Ralf-Peter Jansen showed that She2p co-fractionates with ER membranes independently of Myo4p and She3p, suggesting a novel She2p:ER interaction (Schmid et al., 2006). However, it remains unclear if She2p directly binds to the ER membrane or if this association is mediated by an ER-associated protein or mRNA. These results let assume that localizing mRNAs are co-transported with cortical ER tubules to the bud tip (Figure 5).

1.5.5 Additional *trans*-acting factors involved in *ASH1*-mRNA localization

Besides the She proteins described above, three additional *trans*-acting factors are required for efficient *ASH1* translocation to the bud tip.

INTRODUCTION

Loc1p is an exclusively nuclear protein with enrichment in the nucleolus. It plays roles in ribosome biogenesis as well as in *ASH1*-mRNA localization. So far, Loc1p was mainly described to function in rRNA processing and export of the 60S ribosomal subunit (Harnpicharnchai et al., 2001; Urbinati et al., 2006). However, Loc1p is a double stranded RNA-binding protein, which binds to the *ASH1*-E3 zipcode in the nucleus (Long et al., 2001). The Loc1p:*ASH1* E3 association was recently shown to depend on interaction of Loc1p with She2p, suggesting that She2p recruits Loc1p to the *ASH1* transcript (Shen et al., 2009). Interestingly, both a *LOC1* knockout and exclusion of She2p from the nucleus result in *ASH1*-mRNA delocalization and symmetric distribution of Ash1p in mother and daughter cell (Du et al., 2008; Long et al., 2001; Shen et al., 2009). This suggests that Loc1p may function in translational control of *ASH1* mRNA after recruitment by She2p. The exact mechanism by which nuclear Loc1p controls *ASH1* translation is largely unknown but it was hypothesized that Loc1p might help to generate a specific subtype of ribosomes that are required for the translation of localized mRNAs (Komili et al., 2007). The RNA-binding protein Puf6p functions in translational repression of *ASH1* mRNA (Gu et al., 2004; Shen et al., 2009). Puf6p is a member of the PUF family of RNA-binding proteins and recognizes a conserved PUF protein-recognition sequence (UUGU) in the *ASH1*-E3 zipcode (Gu et al., 2004). In analogy to Loc1p, Puf6p interacts with She2p in the nucleus and is thereby recruited to the *ASH1* mRNA (Shen et al., 2009). Deletion of *PUF6* also results in premature *ASH1* translation and random distribution of Ash1p, demonstrating the role of Puf6 in translational control (Du et al., 2008; Gu et al., 2004; Shen et al., 2009). Furthermore, Puf6p is believed to assemble with the localizing mRNP in the cytoplasm. Here, Puf6p associates with the general translation factor Fun12p/eIF5B and thus inhibits 80S complex assembly (Deng et al., 2008). Upon mRNA localization to the bud tip, casein kinase II (CK2) phosphorylates Puf6p in the N-terminal region, which releases the translational repression. A similar mechanism of translational control also applies to the RNA-binding protein Khd1p (Irie et al., 2002). Khd1p is associated to the localizing mRNP by binding to the *ASH1*-E1 localization element (Irie et al., 2002). Simultaneously, Khd1p interacts with the C-terminal domain of the translation initiation factor eIF4G1, which suppresses translation initiation. At the bud-tip plasma membrane, the casein kinase Yck1p phosphorylates Khd1p, resulting in mRNP disassembly and *ASH1* translation (Paquin et al., 2007). Taken together, the three *trans*-acting factors Loc1p, Puf6p and Khd1p have an important function in mRNA transport by preventing premature translation of the localizing *ASH1* mRNA and thus, ensuring spatially restricted protein synthesis.

1.5.6 Current model of *ASH1*-mRNA localization

The introductory part will conclude with the current model of the mRNA-translocation process in budding yeast (Figure 5). Although at least 24 mRNAs are transported in a She2p-dependent manner, most details are known for the localization of *ASH1* mRNA. Therefore, the model refers to the *ASH1*-localization process but does not exclude the possibility that the localization of other mRNAs is regulated by the same mechanisms. For further details regarding the individual steps and proteins involved, please refer to the figure legend and to chapters 1.5.1 to 1.5.5.

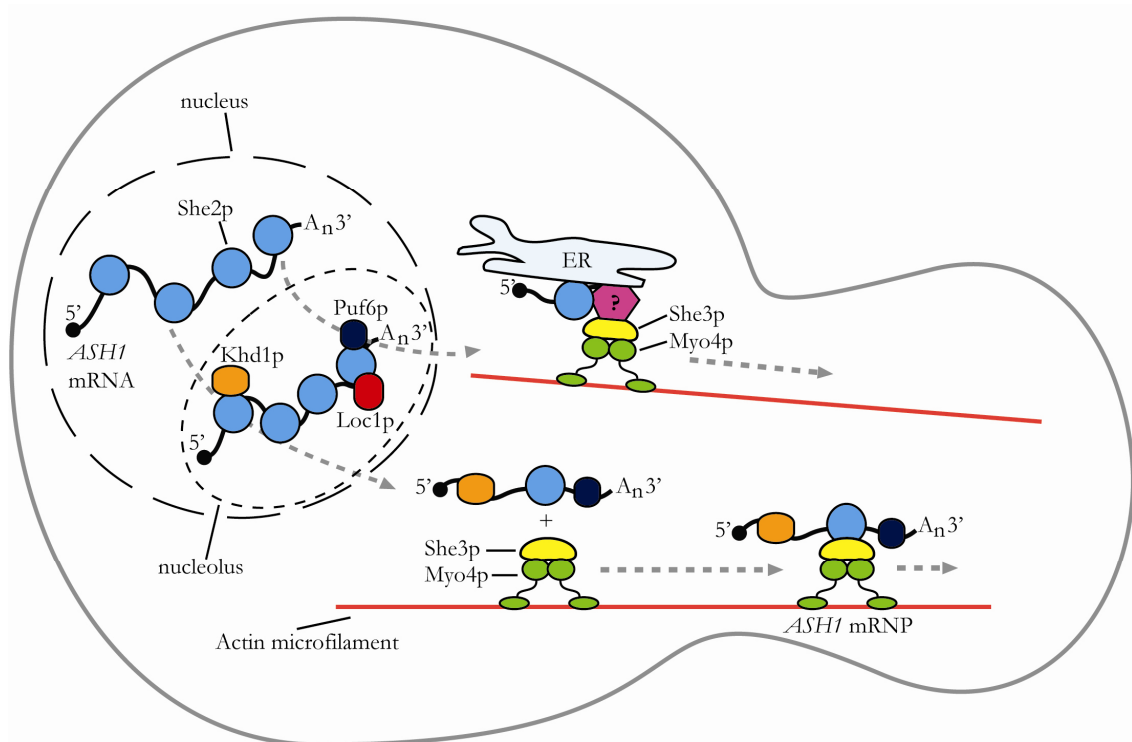


Figure 5: Current model of *ASH1*-mRNA localization. After nuclear import, She2p binds its RNA cargo very likely co-transcriptionally. Then, the She2p:*ASH1* mRNA complex shuttles through the nucleolus, where She2p interacts with Loc1p and Puf6p. Traversing the nucleolus guarantees translational repression of *ASH1* mRNA during the transport process to the bud tip. Following mRNA-dependent export of She2p, the She2p:*ASH1* mRNA complex interacts with the cytoplasmically localized She3p:Myo4p complex and assembles into a functional mRNP. During the Myo4p-dependent transport along the actin cytoskeleton, *ASH1* mRNA is translationally silenced mediated by *ASH1*-associated Khd1p and Puf6p. At the bud tip, the mRNP is anchored at the actin cytoskeleton. The translational repression is released upon phosphorylation of Khd1p by Yck1p and of Puf6p by CK2, respectively, Ash1p is synthesized and represses mating type switching in the daughter cell. Additionally, She3p and Myo4p are required for inheritance of cortical ER. This process is independent of She2p, although She2p alone is able to associate with ER membranes. These results suggest a coordinated movement of localizing mRNAs and ER tubules to the daughter cell.

1.6 Objectives

Localization of mRNA is a widespread mechanism to spatially restrict protein synthesis in a cell. In higher eukaryotes, mRNAs destined for localization are packaged into transport particles of a highly complex architecture. These mRNPs usually comprise dozens of mRNAs and a number of proteins involved in mRNA binding, translational repression, and active transport (Angenstein et al., 2005; Kanai et al., 2004; Ohashi et al., 2002). For correct assembly of such transport complexes, precise and specific interaction of the individual components is indispensable.

In the present study, the requirements for specific transcript binding and mRNP assembly were investigated using the *ASH1* mRNP from *S. cerevisiae* as a model system (Figure 5). In this complex, the RNA-binding protein She2p binds more than 24 mRNAs and connects them via the adapter protein She3p to the motor protein Myo4p. So far, it is unclear how She2p recognizes its target mRNAs and incorporates them into translocation-competent mRNPs. The aim of this study was to identify the requirements for selective mRNA-cargo recognition by She2p in the nucleus and specific mRNP assembly in the cytoplasm. To achieve this, a multidisciplinary approach was chosen. Using a combination of biochemical and structural techniques, the mode of interaction of She2p and its target mRNAs is studied, and based on the crystal structure of She2p (Niessing et al., 2004), novel features required for mRNA binding are investigated. Finally, complementary studies in living yeast cells are used to analyze the role of She2p in mRNP assembly.

2 Results

In this study, I aimed to perform a structural and functional characterization of the assembly of mRNA-transport complexes in *Saccharomyces cerevisiae*. Attempts to obtain structural information about the interaction between She2p and the *ASH1* mRNA at atomic resolution are described in chapter 2.1. The requirements for the assembly of a functional mRNP have been comprehensively analyzed *in vitro* and *in vivo* and are described in chapter 2.2.

2.1 Structural studies on the She2p:RNA interaction

To date, only little structural information is available for individual components of mRNA-transport complexes. The first almost complete atomic model of an integral RNA-binding protein from an mRNA-localization complex could be derived from the crystal structure of She2p from *S. cerevisiae* (see Figure 4 in chapter 1.5.3 and Figure 7 in 2.1.2) (Niessing et al., 2004). However, the mode of interaction with its different RNA targets is largely unknown. Therefore, in the first part of my thesis I aimed at determining the three-dimensional structure of She2p in complex with one of its RNA cargoes by X-ray crystallography. From this, I hoped to obtain mechanistic insights into the She2p:RNA interaction. In a different approach, initial NMR experiments were performed with the intention to map the RNA-binding sites in She2p.

2.1.1 Purification of She2p variants

A number of She2p variants were used for structural studies (described in the present chapter 2.1) and for biochemical analyses (see chapter 2.2). Without exception, all She2p variants were purified to near homogeneity using affinity chromatography, anion-exchange chromatography, and size-exclusion chromatography as described in chapter 4.9.4. A representative elution profile of wild-type She2p from a size-exclusion chromatography column is shown in Figure 6.

RESULTS

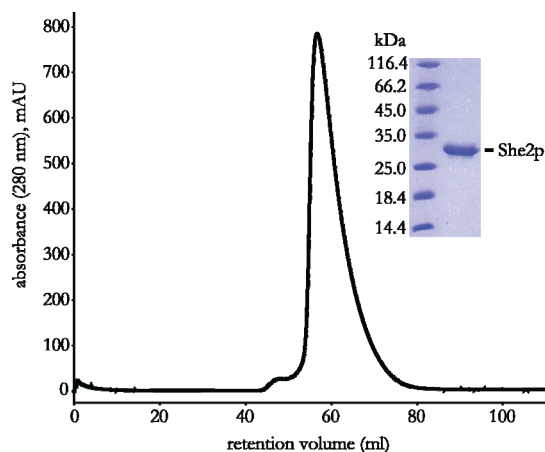


Figure 6: Elution profile of wild-type She2p from a Superdex S200 16/60 size-exclusion column. Absorbance at 280 nm was measured. Wild-type She2p elutes in a single peak at 60 ml retention volume. The inset shows SDS-PAGE analysis of a representative peak fraction.

2.1.2 Designing She2p variants for crystallization

The selection of appropriate She2p variants for crystallization in complex with RNA proved to be crucial. Previous studies revealed that full length-She2p yields only poor-quality crystals and that the four cysteine residues in She2p hamper crystallization (Niessing et al., 2004). Therefore, a mutated She2p fragment had been used for initial crystallization and structure determination of She2p alone (Figure 7) (Niessing et al., 2004). This protein fragment, termed She2p-(6-239)- Δ Cys, lacks five amino acids at the very N-terminus and seven amino acids at the very C-terminus. In addition, all cysteine residues were exchanged by serine. Limited-proteolysis experiments and RNA-binding studies suggested that the N-terminus of She2p is most likely disordered and functionally less relevant, whereas the C-terminus might be important for RNA binding (Niessing et al., 2004). These findings did not permit many variations in designing new She2p constructs for co-crystallization with RNA. Thus, the She2p variants She2p-(6-246)- Δ Cys and She2p-(6-243)- Δ Cys were created and analyzed for their ability to bind *ASH1*-zipcode elements (see chapter 2.1.3.1).

Based on the She2p-crystal structure, two additional She2p variants were created for crystallization experiments, in which flexible or protruding regions were deleted (Figure 7). In general, deletion of unstructured or flexible regions is supposed to facilitate crystallization of molecules and therefore, crystallization of an She2p:RNA complex might be alleviated. The variant She2p- Δ loop lacks residues 81 to 89 in context of full-length She2p. This loop-containing region is located at the bottom of She2p. Since this region was not entirely visible in the electron density, it is most likely flexible (Figure 7) (Niessing et al., 2004). The second variant, She2p- Δ helixE, has deleted amino acids 174 to 183 in context of full-length She2p. This deleted region comprises an α -helix (helix E), which protrudes at right angles out of the

She2p dimer (Figure 7). In the crystal lattice, several crystallographic contacts between She2p dimers could be observed (see also chapter 2.2.17, Figure 29) (Niessing et al., 2004). In one case, crystal contacts formed between two She2p dimers in an upside-down orientation at the helix E-containing side. Only melting of the protruding helices E of both She2p dimers allowed for the formation of these dimer-dimer contacts in the crystal lattice. Deletion of this helix would result in an entirely globular molecule and thus, crystallization of She2p in complex with RNA might be promoted.

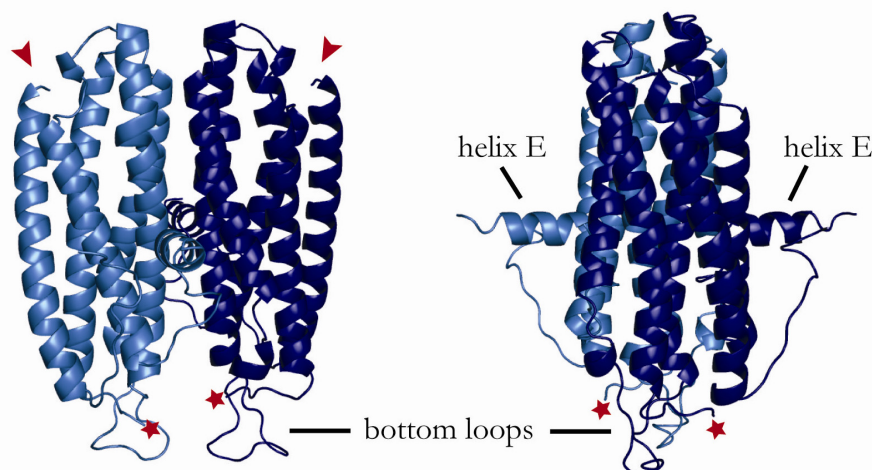


Figure 7: She2p crystal structure. Shown is the crystal structure of She2p (Niessing et al., 2004) with one monomer colored in light blue and the other monomer colored in dark blue. The cartoon shows the She2p dimer in front view (left) and side view (right). The crystal structure lacks five amino acids at the N-terminus and nine amino acids at the C-terminus. The positions of the N-terminus and the C-terminus are labeled by a red asterisk and a red arrow head, respectively. The loops at the bottom region and the protruding helix E, which were deleted for crystallization experiments, are labeled.

2.1.3 Identification of suitable *ASH1*-RNA fragments for co-crystallization

ASH1 mRNA, the best-characterized RNA target of She2p, contains four zipcode elements (E1, E2A, E2B, and E3), which fold into distinct-stem loop structures (Figure A1 in the appendix) and are specifically recognized by She2p (Böhl et al., 2000; Jambhekar et al., 2005; Long et al., 2000; Olivier et al., 2005). The affinity of wild-type She2p to the *ASH1*-E3 zipcode was previously determined to be $210 \text{ nM} \pm 40 \text{ nM}$ (Niessing et al., 2004). For co-crystallization experiments however, it was indispensable to characterize the She2p:*ASH1* interaction more comprehensively. By this means, different shorter *ASH1*-mRNA fragments could be identified that are strongly bound by She2p and might therefore be suitable for crystallization. A detailed analysis of RNA binding by She2p, including all necessary controls, is presented in chapter 2.2.

2.1.3.1 She2p-(6-246)- Δ Cys is suitable for crystallization with *ASH1*-E3 RNA and minimized *ASH1* zipcodes

At first, the equilibrium-dissociation constants (K_D) of the variants She2p-(6-246)- Δ Cys and She2p-(6-243)- Δ Cys to the *ASH1*-E3 zipcode were determined and compared with the affinity of wild-type She2p (Table 2). The interaction between wild-type She2p and RNA has been investigated in detail and will be described in chapter 2.2. The E3 zipcode was bound strongly by the She2p variant comprising the complete C-terminus (She2p-(6-246)- Δ Cys; $K_D=173 \text{ nM} \pm 12 \text{ nM}$). However, E3 was bound ten-fold weaker by the She2p variant that lacks three amino at the C-terminus (She2p-(6-243)- Δ Cys; Table 2). This finding suggests that the complete C-terminus of She2p is required for proper recognition of the *ASH1*-E3 zipcode and therefore, only She2p-(6-246)- Δ Cys was used for co-crystallization experiments (chapter 2.1.4).

It proved difficult to produce *ASH1* RNAs by *in vitro* transcription in such a high amount and quality needed for crystallization. The length of the *ASH1* zipcodes ranges from 58 to 118 nucleotides. Since the maximum length for reliable chemical synthesis of RNA is about 50 nucleotides, *ASH1* zipcodes cannot be obtained by chemical synthesis. Previously, the *ASH1*-E1 and *ASH1*-E2B zipcodes were minimized to a length of 49 nucleotides (E1min) and 36 nucleotides (E2Bmin), respectively (Jambhekar et al., 2005; Olivier et al., 2005). These shortened zipcodes contain all features that are required for efficient *ASH1*-mRNA localization *in vivo*.

Subsequently, She2p's affinity to the minimized *ASH1* zipcodes was determined. Wild-type She2p and She2p-(6-246)- Δ Cys bound to the E1min RNA with binding constants of $K_D=456 \text{ nM} \pm 78 \text{ nM}$ and $K_D=599 \text{ nM} \pm 52 \text{ nM}$, respectively (Table 2). A comparably higher affinity of both proteins to the E2Bmin zipcode was observed ($K_D=103 \text{ nM} \pm 38 \text{ nM}$ and $K_D=174 \text{ nM} \pm 12 \text{ nM}$, respectively; Table 2). Again, She2p-(6-243)- Δ Cys bound significantly weaker to minimized zipcodes than She2p-(6-246)- Δ Cys (Table 2). I also narrowed down the *ASH1*-E2B zipcode to a length of 21 nucleotides, comprising only the upper part of the stem loop (appendix, Figure A1). This RNA, termed E2Bmin-21, was bound by She2p-(6-246)- Δ Cys with considerable affinity ($K_D=259 \text{ nM} \pm 60 \text{ nM}$; Table 2).

Taken together, the She2p variant She2p-(6-246)- Δ Cys was selected for co-crystallization studies with the RNA oligonucleotides *ASH1*-E3, E1min, E2Bmin, and E2Bmin-21.

RESULTS

RNA	K_D wild-type She2p	K_D She2p-(6-246)- Δ Cys	K_D She2p-(6-243)- Δ Cys	K_D She2p- Δ loop	K_D She2p- Δ helixE
<i>ASH1</i> E1	236 nM \pm 14	n.d.	n.d.	n.d.	2.3 μ M \pm 0.2
<i>ASH1</i> E2A	258 nM \pm 46	n.d.	n.d.	n.d.	0.94 μ M \pm 0.15
<i>ASH1</i> E2B	308 nM \pm 3	n.d.	n.d.	n.d.	6.8 μ M \pm 2.1
<i>ASH1</i> E3	102 nM \pm 16	173 nM \pm 12	1.7 μ M \pm 0.5	129 nM \pm 12	0.76 μ M \pm 0.06
E1min	456 nM \pm 78	599 nM \pm 52	942 nM \pm 253	630 nM \pm 121	n.d.
E2Bmin	103 nM \pm 38	174 nM \pm 12	461 nM \pm 92	144 nM \pm 9	n.d.
E2Bmin-21	n.d.	259 nM \pm 60	150 nM \pm 31	2.56 μ M \pm 0.9	n.d.

Table 2: RNA-binding studies with She2p-wild type and She2p variants. The table summarizes the equilibrium-dissociation constants (K_D) obtained with indicated She2p variants. The label “n.d.” indicates that this experiment was not performed. For further details, please refer to main text. A detailed analysis of RNA binding by She2p, including control experiments, is shown in chapter 2.2.

2.1.3.2 Binding studies with She2p- Δ loop and She2p- Δ helixE

The variant She2p- Δ loop bound to the *ASH1*-E2Bmin RNA with near wild-type affinity only (K_D =144 nM \pm 9 nM; Table 2). Thus, only the E2Bmin oligonucleotide was chosen for co-crystallization experiments.

She2p- Δ helixE showed a significantly impaired affinity to all *ASH1*-localization elements compared to wild-type She2p (Table 2), rendering this variant inappropriate for crystallization experiments in complex with RNA. A detailed investigation of RNA binding by She2p- Δ helixE is described in chapter 2.2.

2.1.4 Co-crystallization experiments with *ASH1*-E3, E1min, and E2Bmin RNA

She2p-(6-246)- Δ Cys and She2p- Δ loop were each concentrated to 4 mg/ml and incubated with a 1.5 molar excess of RNA. Since the stoichiometry of She2p and different *ASH1* zipcodes is not completely clarified, a 1:1 stoichiometry (one RNA zipcode per She2p monomer) was assumed and a molar excess of RNA was used in the co-crystallization experiments. Initial crystallization screens were set up in 96-well plates with 1 μ l drops consisting of 0.5 μ l protein:RNA solution and 0.5 μ l precipitant solution. A number of different commercially available 96-well initial screens were tested, but unfortunately no crystals were obtained with She2p-(6-246)- Δ Cys in complex with *ASH1*-E3, E1min, or E2Bmin RNA, respectively. Small

RESULTS

crystals of the She2p- Δ loop:E2Bmin complex were obtained in JBScreen Classic-I, #D12 (20 % (w/v) PEG 4000, 20 % 2-propanol, 100 mM sodium citrate) and could be refined to a maximum size of 35 μ m x 20 μ m x 10 μ m in 20 % (w/v) PEG 4000, 14 % 2-propanol, 100 mM sodium citrate. Unfortunately, crystals were not uniformly shaped, mostly two-dimensional, and did not diffract X-rays when tested at DESY (Hamburg).

2.1.5 Crystallization of She2p in complex with *ASH1* E2Bmin-21 variants

Crystallization trials with She2p-(6-246)- Δ Cys in complex with E2Bmin-21 RNA were performed by mixing 0.25 μ l of protein solution (4 mg/ml) with 0.25 μ l of E2Bmin-21 RNA (1.5-molar excess) and 0.5 μ l precipitant solution in 96-well plates. Microcrystals and crystals of irregular shape were obtained in a few initial conditions, which mostly contained PEG. Microcrystals grown in JBScreen Classic-I, # F8 (30 % (w/v) PEG 6000, 1 M lithium chloride, 100 mM sodium acetate) were subjected to extensive refinement of pH and the PEG/salt concentration, as well as by adding different salts or glycerol. Finally, larger crystals of an irregular shape could be obtained with a maximum size of 30 μ m x 20 μ m x 20 μ m. But still, the crystal quality was insufficient for diffracting X-rays when tested at several beamlines at DESY (Hamburg) and at the ESRF (Grenoble).

When She2p or RNA was omitted from the crystallization set-up no crystals were obtained, indicating that in the JBS-I-F8 condition the She2p:E2Bmin-21-RNA co-complex was crystallized. Therefore, the E2Bmin-21-RNA sequence was modified in order to increase crystal quality. These modifications included the generation of 3' or 5' overhangs, variations in the length of the stem, or exchanging base-paired nucleotides in the stem. In total, nine additional RNAs (termed E2Bshort 01 to 09, see Table A1 in the appendix) were generated based on the E2Bmin-21 RNA and used in crystallization experiments with She2p-(6-246)- Δ Cys. Three-dimensional crystals could be obtained with three different RNAs (E2Bshort-01, E2Bshort-04, E2Bshort-05; Figure 8). Interestingly, crystals of the different complexes grew in very similar conditions, but their quality and their shape were completely different. In no case, crystals were obtained when protein or RNA was omitted from the crystallization experiment, indicating that She2p:RNA co-complexes were crystallized. Crystals of the She2p:E2Bshort-04 complex were grown to small plates and showed no diffraction when tested at the ESRF (Figure 8 a). Elliptic crystals with a maximum size of 40 μ m x 15 μ m x 15 μ m were obtained for the She2p:E2Bshort-05 complex (Figure 8 b). At best, these crystals diffracted to 10 Å

RESULTS

resolution at different beamlines at the ESRF. The best crystals were obtained in complex with the E2Bshort-01 RNA (Figure 8 c). The maximum size of the cubic crystals never exceeded $35\ \mu\text{m} \times 25\ \mu\text{m} \times 25\ \mu\text{m}$. When tested at the ESRF, they diffracted to a maximum resolution of only $9\ \text{\AA}$. Considerable effort was made to improve the diffraction quality of the crystals. This included further optimization of the growth condition to increase the crystal size, the use of additive screens, soaking the crystals with different heavy-metal compounds, shifting the temperature to $4\ ^\circ\text{C}$, using different cryoprotectants, and improving the crystal freezing technique. However, none of these attempts resulted in better diffraction of the crystals and no data sets could be collected.

In summary, diffracting crystals were obtained for She2p-(6-246)- ΔCys in complex with two different RNA variants of the *ASH1*-E2B zipcode. Unfortunately, the crystal quality was not high enough to record any data set and therefore, no atomic model of She2p in complex with an RNA target could be derived.

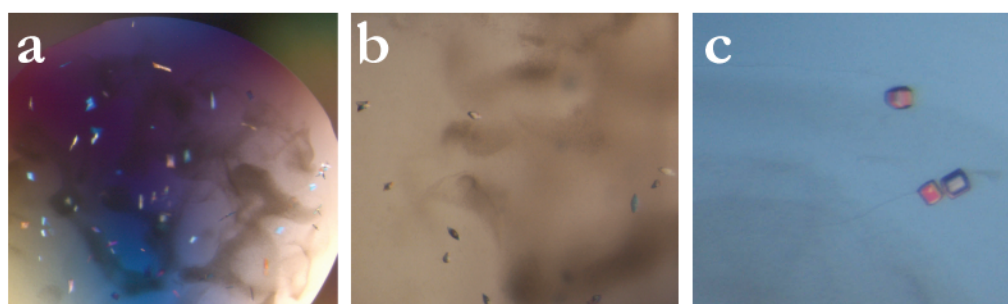


Figure 8: Crystals of She2p:RNA co-complexes. **a)** Crystals of She2p-(6-246)- ΔCys in complex with the E2Bshort-04 RNA were obtained in 26 % (w/v) PEG 6000, 0.7 M LiCl_2 , 100 mM sodium acetate. **b)** Elliptic crystals of the She2p-(6-246)- ΔCys : E2Bshort-05 co-complex were grown in 22 % (w/v) PEG 6000, 0.8 M LiCl_2 , 100 mM sodium acetate. **c)** Cubic crystals were obtained for the She2p-(6-246)- ΔCys : E2Bshort-01 co-complex in 24 % (w/v) PEG 6000, 0.5 M LiCl_2 , 100 mM sodium acetate. At best, these crystals diffracted X-rays to a maximum resolution of $9\ \text{\AA}$.

2.1.6 Can She2p crystals be soaked with RNA?

Another idea for obtaining a crystal structure of She2p in complex with RNA was to reproduce the original She2p crystals (Niessing et al., 2004) and to soak them with short RNA oligonucleotides. The previously identified conserved sequence elements of 3 to 9 bases in *ASH1* zipcodes (Jambhekar et al., 2005; Olivier et al., 2005) were used as target RNAs, because they should be small enough to diffuse into the crystal lattice. First, the affinity of

RESULTS

She2p-(6-239)- Δ Cys (see chapter 2.1.2) to short biotinylated RNA oligonucleotides was measured by Surface Plasmon Resonance (SPR) using a Biacore 3000 system. Surprisingly, She2p-(6-239)- Δ Cys was not able to bind to any of those RNA sequences in contrast to wild-type She2p, which bound with equilibrium-dissociation constants of 18.6 μ M and 20.7 μ M to the 9-mer (CGACGAAAA) and the 6-mer sequences (CGACGA), respectively (see chapter 2.2.4). The observed weak affinities imply that for efficient RNA binding a minimal length and most likely secondary-structure elements are necessary. For that reason, the planned soaking experiments were not performed.

2.1.7 Exploratory NMR experiments with She2p

Nuclear Magnetic Resonance (NMR) data were acquired and analyzed by Dr. Bernd Simon from the laboratory of Prof. Michael Sattler (EMBL Heidelberg/Helmholtz Zentrum München).

In parallel to crystallization experiments, NMR experiments were performed with She2p with the intention to map the RNA-binding region in She2p. Therefore it was planned to perform NMR-titration experiments with the *ASH1*-E2Bmin zipcode. This RNA was chosen because of its strong affinity to She2p, its relatively small size, and its availability in high amount by chemical synthesis. Significant binding of RNA to She2p would result in chemical-shift perturbations in a two-dimensional ^1H - ^{15}N -HSQC spectrum. From the titration data, I planned to calculate the equilibrium dissociation constant of the She2p:RNA interaction and to identify residues interacting with *ASH1*-E2Bmin RNA. Unfortunately, the 2D ^1H - ^{15}N -HSQC spectrum recorded for ^{15}N -labeled wild-type She2p resulted in only a few peaks rendering She2p inappropriate for NMR measurements (appendix, Figure A2). This result is most likely due to the relatively large molecular weight of the She2p oligomer, which was determined by me in later experiments (see chapters 2.2.14 to 2.2.18). Thus, the planned RNA-titration experiments with *ASH1*-E2Bmin RNA could not be performed by NMR.

2.2 Functional analysis of mRNP assembly

The RNA-binding protein She2p couples the mRNA cargo to the motor Myo4p via its interaction with the adapter protein She3p. For that reason, She2p is believed to be responsible for the selection of the correct RNA cargoes and their incorporation into the transport particles. However, the mechanistic details underlying this selection process are still unknown. Therefore, in the second part of my thesis, the requirements for the assembly of a functional *ASH1* mRNP and for *ASH1*-mRNA localization in *S. cerevisiae* have been investigated.

2.2.1 She2p binds to zipcode elements with nanomolar affinity

In *S. cerevisiae*, more than 30 mRNAs are localized in a She2p-dependent manner, but only for a small subset of these RNAs the zipcode elements are identified. Surprisingly, only She2p's affinity to the *ASH1*-E3 zipcode has been quantified to date (Niessing et al., 2004), whereas the affinity to the other zipcodes is unknown. I performed RNA filter-binding assays in order to quantify the binding of wild-type She2p to all four *ASH1* zipcodes (E1, E2A, E2B, E3), to the zipcode of the *EAR1* mRNA, and to the N-terminal zipcode of the *WSC2* mRNA (*WSC2N*) (Figure 9). Before the experiments, RNA integrity was always checked in native polyacrylamide gels (not shown). All She2p variants used for biochemical and biophysical analysis were purified as described in chapter 2.1.1.

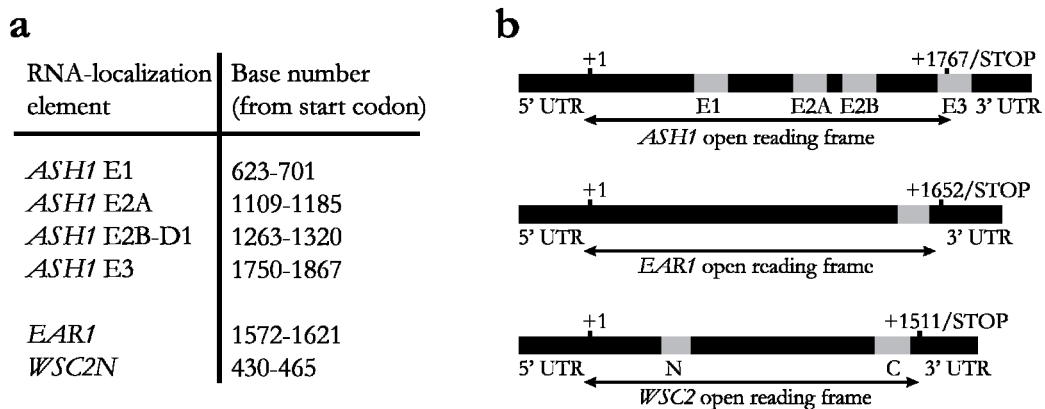


Figure 9: Illustration of RNA-localization elements used in this study. The table in **a)** lists the position of the different zipcodes in the respective mRNAs relative to the start codon. Schematic drawing in **b)** represents the location of the zipcodes in the respective mRNA as listed in **a)**. The exact lengths of the 5' and 3' untranslated regions (UTR) are unknown and therefore the UTRs are displayed with an arbitrary size.

RESULTS

Secondary-structure predictions revealed that these zipcodes fold into specific stem-loop structures (Chartrand et al., 2002; Gonzalez et al., 1999; Jambhekar et al., 2005; Olivier et al., 2005). The secondary structures of all RNAs used in this study were predicted with the MC-Fold algorithm (Parisien and Major, 2008) and are shown in Figure A1 in the appendix. In order to measure RNA binding, serial dilutions of the respective protein (all protein concentrations used for RNA-binding studies are listed in chapter 4.11.4) were incubated with a constant amount of radiolabeled RNA for 20 minutes at 20 °C and then blotted onto a nitrocellulose membrane using a 96-well microfiltration apparatus. The nitrocellulose membrane binds only protein and protein:RNA complexes, respectively, whereas unbound RNA is not retained. Protein-bound RNA was quantified by phosphoimaging and equilibrium-dissociation constants (K_D) were calculated using the Langmuir isotherm. The binding assays showed that wild-type She2p binds to all zipcode elements with affinities between 0.10 μ M and 0.77 μ M (Figure 10 a). Strongest binding was observed to the *ASH1*-E3 element with a K_D =0.10 μ M \pm 0.02 μ M (Figure 10 a, b), which was an even stronger association than described previously (Niessing et al., 2004). The weakest binding was calculated for the interaction of She2p with the *EAR1* zipcode (K_D =0.77 μ M \pm 0.24 μ M) (Figure 10 a). Taken together, the binding constants measured with wild-type She2p demonstrate tight binding to zipcode elements of three different bud-localizing mRNAs from *S. cerevisiae*.

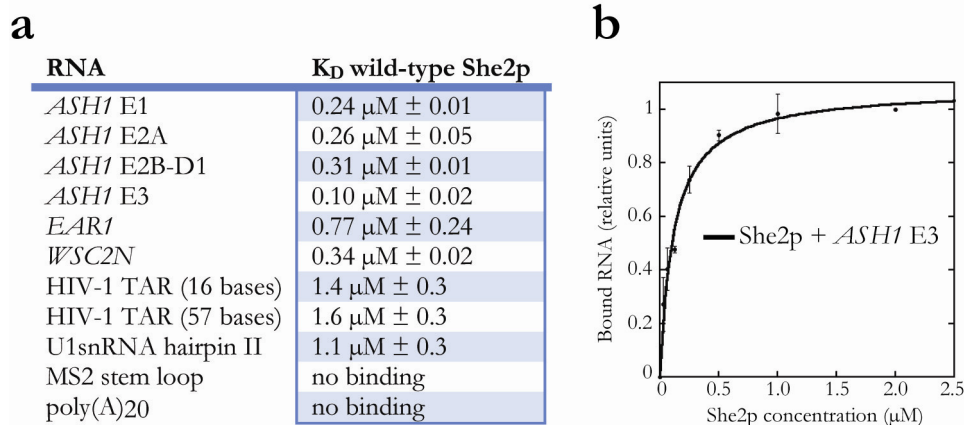


Figure 10: RNA-binding studies with wild-type She2p. The table in **a)** summarizes the results of RNA filter-binding experiments performed with wild-type She2p. She2p binds to zipcodes of bud-localized mRNAs with equilibrium-dissociation constants between 0.1 μ M (*ASH1* E3) and 0.77 μ M (*EAR1*). Binding to the unrelated stem-loop forming RNAs HIV-1 TAR and U1snRNA hairpin II is about two- to ten-fold weaker, whereas the unrelated MS2-stem loop and single stranded poly(A) RNA are not bound by She2p. Shown is the mean of at least three independent experiments, \pm shows the standard deviation. The label “no binding” indicates that at the highest protein concentration measured, no significant binding was observed (MS2 stem loop: 6 μ M; poly(A)₂₀: 32 μ M). **b)** depicts a representative curve fit of RNA-binding experiments performed with wild-type She2p and the *ASH1*-E3 zipcode. The curve was fitted using the Langmuir isotherm.

2.2.2 She2p has considerable affinity to stem-loop containing RNAs

Control experiments with a linear, unstructured RNA consisting of 20 adenine nucleotides (poly(A)₂₀) revealed no detectable She2p binding (Figure 10 a). However, wild-type She2p bound to the unrelated stem loops of HIV-1 TAR RNA (16 nucleotides) and U1snRNA (22 nucleotides) with a $K_D=1.4 \mu\text{M} \pm 0.3 \mu\text{M}$ and a $K_D=1.1 \mu\text{M} \pm 0.3 \mu\text{M}$, respectively (Figure 10 a). These affinities are only two- to ten-fold weaker than the affinity to the specific zipcode RNAs. A longer stem loop of HIV-1 TAR comprising 57 nucleotides was bound with a $K_D=1.6 \mu\text{M} \pm 0.3 \mu\text{M}$, indicating that the length of the unrelated stem-loop containing RNA has no significant influence on She2p binding. In contrast, wild-type She2p failed to bind to the MS2 hairpin, another unrelated and thus non-specific stem-loop RNA, at concentrations up to 6 μM (Figure 10 a). Taken together, these findings suggest that She2p is able to bind non-specifically to a subset of unrelated RNAs, which fold into stem-loop structures.

2.2.3 She2p monomers do not bind to RNA

A previous study has shown that She2p needs to dimerize for RNA binding (Niessing et al., 2004). Mutation of amino acids that are located in the dimer interface of She2p abolished *ASH1*-E3 binding *in vitro* (C68Y and S120Y, respectively) and *ASH1* localization *in vivo* (S120Y). Here, I extended the binding studies with the variant She2p-S120Y (Figure 11 a, b) and analyzed the binding to all *ASH1* zipcodes E1, E2A, E2B, and E3, as well as to the unrelated stem-loop containing RNAs HIV-1 TAR (16 bases) and U1snRNA hairpin II. Using filter-binding experiments, no binding of She2p-S120Y to any of those RNAs could be observed at a concentration up to 12 μM and no binding constants could be calculated (Figure 11 c). These findings are consistent with the previous study (Niessing et al., 2004) and indicate that She2p dimers are required for binding to both specific target RNA (*ASH1*) and non-specific RNA-stem loops (HIV-1 TAR, U1snRNA). Furthermore, it can be assumed that *ASH1*-mRNA localization *in vivo* completely depends on oligomeric She2p.

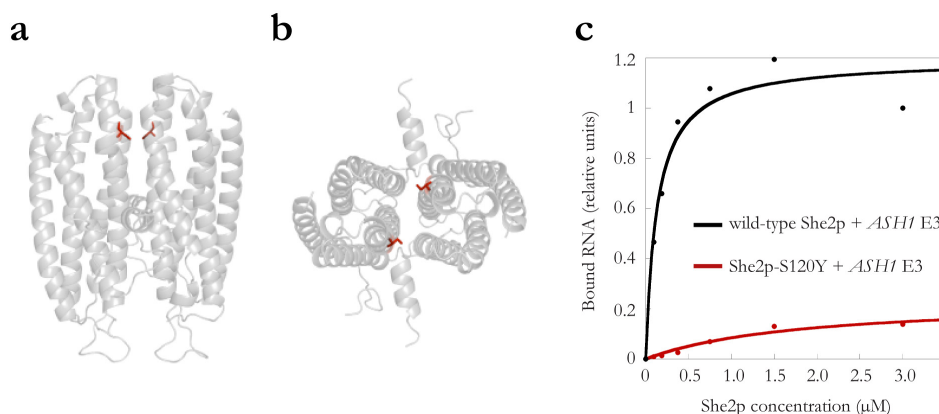


Figure 11: She2p dimerization is indispensable for RNA binding. The cartoons in **a)** and **b)** highlight in red serine 120 in the dimer interface that was mutated to tyrosine. **a)** shows She2p from the front and **b)** from the top. **c)** The mutant protein She2p-S120Y neither binds to any *ASH1* zipcode nor to the stem-loop containing RNAs HIV-1 TAR and U1snRNA. The graph shows a representative plot of the relative bound E3-RNA fraction by She2p-S120Y compared to the RNA fraction that is bound by wild-type She2p. No dissociation constant of She2p-S120Y to any RNA could be determined (maximum protein concentration measured: 12 μM).

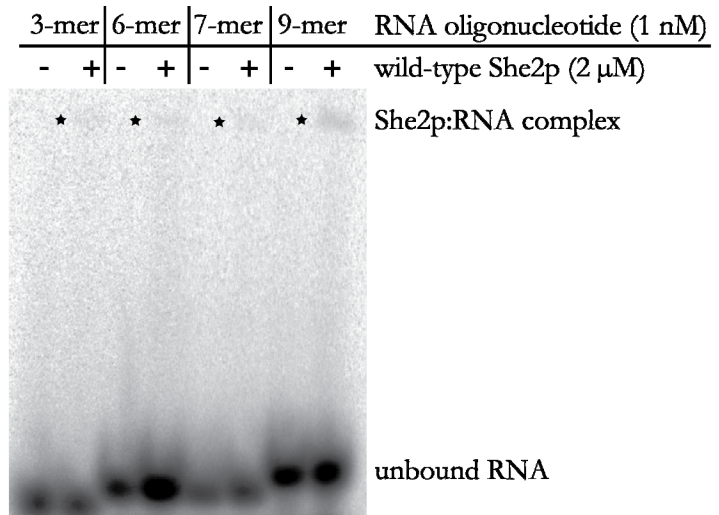
2.2.4 She2p has low affinity to a short zipcode-consensus sequence

The observation that She2p can bind to unrelated stem-loop containing RNAs with significant affinity raised the question how She2p selectively recognizes its mRNA targets in the nucleus. Stem-loops are abundant secondary structures in any kind of RNAs (Svoboda and Di Cara, 2006). Thus, She2p could in principle bind to a number of mRNAs, tRNAs, or rRNAs in the nucleus, translocate them to the cytoplasm and incorporate them into the translocation complex. The only modest difference in affinity of She2p to specific and non-specific RNAs is unlikely to explain how She2p discriminates between bud-localizing mRNA and non-specific RNA.

One possibility is that She2p selects bud-localizing RNAs by recognizing conserved sequence motifs of 3 to 9 nucleotides in length within the zipcode, as it was recently proposed by Jambhekar et al. 2005 and Olivier et al., 2005. To test if She2p interacts with these short motifs alone, electrophoretic mobility shift assays (EMSA) were performed. Wild-type She2p was incubated with the respective radiolabeled RNA oligonucleotide and the RNA:protein complexes were separated from unbound RNA in a native polyacrylamide gel. As shown in Figure 12, She2p was able to bind to each of the short sequence elements, but the interaction was barely visible in EMSA experiments. Furthermore, a significant amount of unbound RNA was noticed, although a 2000-fold molar excess of protein was used in the individual reactions. This observation indicates that the monitored She2p:RNA association might be a rather transient interaction.

RESULTS

Figure 12: She2p binding to short RNA-consensus sequences can be detected. Electrophoretic mobility shift assays using a native 7 % polyacrylamide gel revealed binding of wild-type She2p to a 3-mer, 6-mer, 7-mer, and 9-mer RNA oligonucleotide (indicated by an asterisk). In the experiment, a 2000-fold molar excess of protein was used. However, a large amount of unbound RNA was detected, which might reflect a transient interaction.



In order to quantify the identified interactions and to analyze their stability, I performed SPR experiments using a Biacore 3000 system (see also chapter 2.1.6). RNA oligonucleotides (9-mer, 6-mer, 3-mer) with a 5' biotin tag were immobilized on a streptavidin-coated chip surface and She2p was floated over the surface. If any interaction occurs, an increase in the response signal can be monitored. Analysis of the obtained sensorgrams revealed a fast association and dissociation of the complex (Figure 13 a). From the measured response curves no on-rates (K_{on}) or off-rates (K_{off}) could be derived in order to calculate the equilibrium-dissociation constant K_D ($K_D = K_{off} / K_{on}$). However, the binding constant was calculated from steady state response signals of multiple injections with increasing protein concentrations by plotting the response signals at the equilibrium against the respective protein concentration. The experiments revealed that She2p bound to the 9-mer motif and the 6-mer motif with a binding constant of 18.6 μ M and 20.7 μ M, respectively (Figure 13 a-c). In contrast, no detectable binding was observed to the 3-mer motif (CGA) up to a protein concentration of 200 μ M (Figure 13 c). I cannot exclude the possibility that the 3-mer RNA failed to bind due to its close proximity to the dextran matrix of the chip. In comparison to She2p's affinity to "full-length" zipcodes (chapter 2.2.1), the binding constants measured to these short RNA sequences are significantly weaker. These findings indicate that short RNA motifs alone are unable to mediate specific She2p binding. However, it cannot be excluded that they act in combination with other nucleic acid features like secondary- and tertiary-structure elements to achieve specific zipcode recognition by She2p.

To test whether the combination of secondary-structure elements and an RNA-consensus motif results in a zipcode element, a rather simple approach was chosen: nucleotides in the

RESULTS

upper region of the HIV-1 TAR stem loop were replaced by the conserved 6-mer motif (CGACGA) in a way that preserved the predicted secondary structure. If the RNA motif would function in combination with a simple secondary structure, an affinity would be expected that is comparable to the affinity to zipcodes. As a control, the RNA motif was inserted in antisense orientation (GCUGCA), for which no increase in affinity would be expected. Secondary-structure predictions of both “artificial” zipcodes are shown in Figure A1 in the appendix. RNA filter-binding experiments revealed that wild-type She2p binds with similar affinity to both “artificial” zipcodes ($K_D=1.2 \mu\text{M} \pm 0.2 \mu\text{M}$ and $K_D=1.0 \mu\text{M} \pm 0.2 \mu\text{M}$, respectively) as to the “wild-type” HIV-1 TAR RNA (compare Figure 13 c with Figure 10 a). This finding indicates that the nature of a zipcode is much more complex and cannot be simply generated by combining a conserved RNA motif with secondary-structure elements.

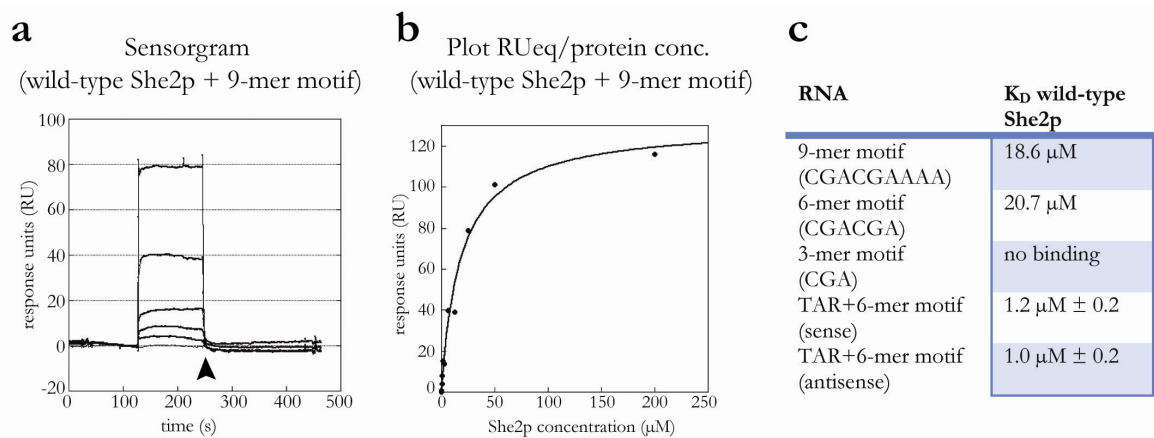


Figure 13: She2p has low affinity to short RNA motifs in zipcode elements. **a)** shows the response signals (in response units) of the She2p:9-mer RNA interaction obtained by SPR experiments. The response curves represent multiple experiments with increasing protein concentrations. Please note the rapid dissociation of the protein:RNA complex when protein injection to the chip surface was stopped (marked by an arrow head). **b)** Plot of the response units at equilibrium against the respective She2p concentration. The binding constant was calculated by fitting the curve using the Langmuir isotherm. The table in **c)** summarizes the results of SPR experiments and filter-binding experiments performed with wild-type She2p and short RNA motifs and modified TAR-stem loops. For She2p’s interaction with the 9-mer and the 6-mer motif, a K_D of 18.6 μM and 20.7 μM was calculated, respectively. No affinity could be measured to the 3-mer motif up to protein concentration of 200 μM . Insertion of the 6-mer motif in any orientation into the HIV-1 TAR-stem loop did not result in increased affinity of She2p compared to the binding to the unmodified TAR-stem loop ($K_D=1.4 \mu\text{M} \pm 0.3$; Figure 10 a).

2.2.5 Identification of conserved regions in She2p for mutational analysis

Besides features on the RNA level, specificity to bud-localizing mRNAs must also be achieved by She2p itself. To test the requirements of She2p for mRNA localization, conserved regions in She2p that are potentially involved in RNA binding were selected for mutational analysis

RESULTS

(Figure 14). In a previous study the basic helical hairpin motif was already described as an unusual RNA-binding motif in She2p (Niessing et al., 2004). In addition, mutation of threonine 47 and leucine 130 in the upper surface region of She2p resulted in reduced affinity to *ASH1* E3 (Niessing et al., 2004). Both regions are highly conserved throughout other yeast species (Figure 14 a, b and Figure A3 in the appendix), suggesting that other conserved regions in She2p might be functionally relevant as well. Two additional conserved surface features could be identified in She2p. The small helix E that protrudes at right angles out of the globular helical bundle and the very C-terminus of She2p both show a high degree of sequence conservation (Figure 14 c, d and Figure A3 in the appendix). The C-terminal nine amino acids of She2p were not visible in the crystal structure, most likely due to their flexibility. All four regions in She2p mentioned above were mutated (Figure 14 e-h) and the respective mutant proteins were tested for their binding to zipcode RNAs and to unrelated stem-loop containing RNAs.

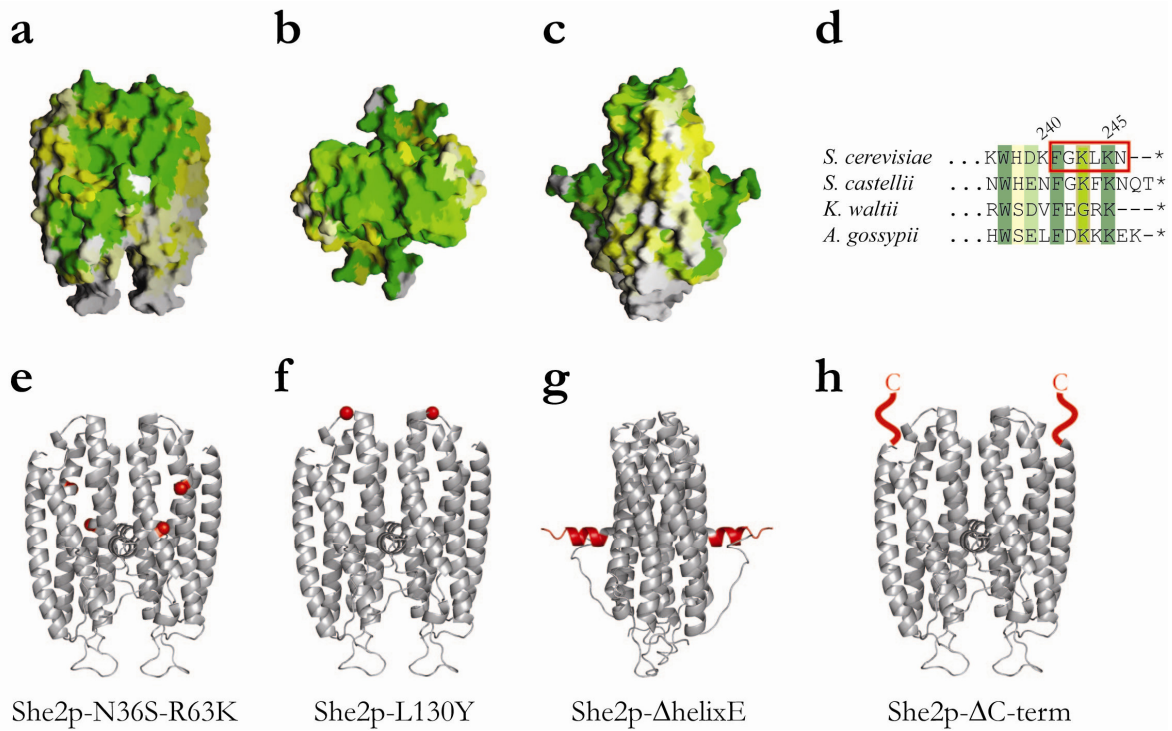


Figure 14: Identification of conserved surface regions in She2p. **a)-c)** The sequence conservation of She2p was plotted onto the surface of the She2p X-ray crystallographic model. Cartoons show She2p in front view (a), top view (b), and side view (c). **d)** shows a close-up of the C-terminal region from the She2p-sequence alignment (Figure A3 in the appendix). Residues deleted for mutational analysis are marked by a red box. The color code ranging from yellow to green represents 40 % to 100 % sequence identity. **e)-h)** show cartoons of She2p mutants, which were created according to the analysis of sequence conservation. Mutated regions are labeled in red and represent the mutation of N36 and R63 in the basic helical hairpin (e), mutation of L130 (f), deletion of helix E (g), and deletion of the C-terminal 6 amino acids (h).

2.2.6 The basic helical hairpin motif provides a general binding platform for stem-loop containing RNA

In order to completely abolish RNA binding by the basic helical hairpin motif, asparagine 36 was mutated to serine and arginine 63 was mutated to lysine, resulting in the double mutant She2p-N36S-R63K (Figure 14 a, e and Figure 15 a). Mutation of both residues was already described to impair *ASH1*-mRNA binding *in vitro* and *ASH1*-mRNA localization *in vivo* (Gonsalvez et al., 2003; Niessing et al., 2004). Binding of the mutant protein to the *ASH1* zipcodes, the *EAR1* zipcode, and to the unrelated RNAs HIV-1 TAR and U1snRNA hairpin II, respectively, was assessed by filter-binding experiments. She2p-N36S-R63K completely failed to bind to any RNA, regardless if target or non-target RNA (Figure 15 b, c). This result suggests that the positively charged basic helical hairpin motif acts as general RNA-binding platform. This finding is further supported by the observation that the She2p-N36S-R63K mutant has completely lost its ability to localize *ASH1* mRNA *in vivo* (Du et al., 2008).

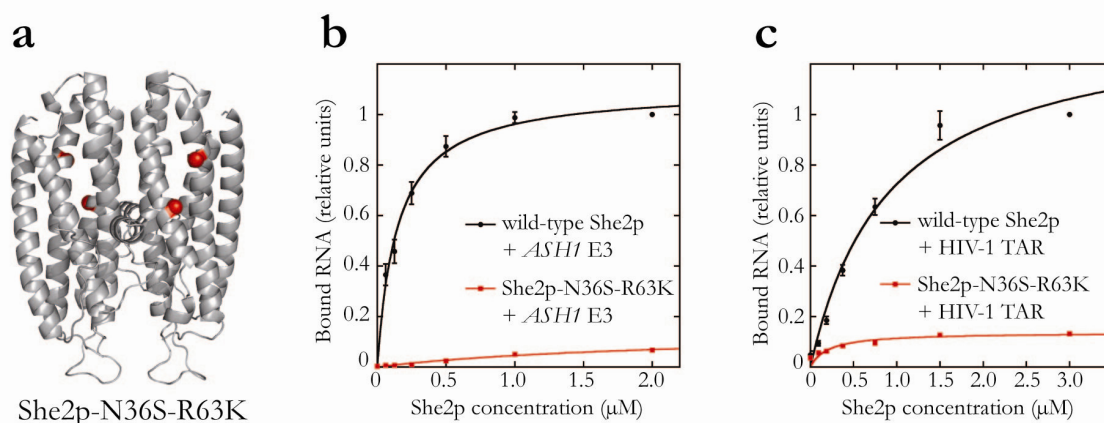


Figure 15: The basic helical hairpin motif provides a binding platform for stem-loop containing RNA. The cartoon in **a)** highlights in red asparagine 36 and arginine 63 in the basic helical hairpin motif that were mutated to serine and lysine, respectively. **b)** and **c)** The mutant protein She2p-N36S-R63K has affinity neither to zipcode-containing RNAs nor to the stem-loop containing RNAs HIV-1 TAR and U1snRNA. The graphs show the relative bound RNA fraction by She2p-N36S-R63K compared to the RNA fraction that is bound by wild-type She2p. A dissociation constant of She2p-N36S-R63K to neither RNA could be determined (maximum protein concentration measured: 32 μM).

2.2.7 The upper uncharged surface region is also required for recognition of stem-loop containing RNAs

Mutation of leucine at position 130 was described before to reduce She2p's affinity to the *ASH1*-E3 zipcode and to impair the She2p:She3p interaction *in vivo* (Gonsalvez et al., 2003; Niessing et al., 2004). However, no binding constants were calculated. Therefore, leucine 130

RESULTS

was mutated to tyrosine (Figure 14 b, f and Figure 16 a) and the binding of the mutant She2p-L130Y to zipcode-containing RNAs and unrelated stem-loop RNAs was assessed using filter-binding experiments (Figure 16 b-d). A significantly reduced affinity of She2p-L130Y to all zipcode elements was observed, ranging from a 2-fold reduction (*EAR1*) to an almost 30-fold reduction (*WSC2N*) (Figure 16 b). No binding could be measured to the unrelated stem-loops of the HIV-1 TAR RNA and the U1snRNA up (Figure 16 b-d). These findings show that mutation of a single hydrophobic residue in the upper surface region of She2p drastically impairs the binding to target and non-target RNAs, respectively. Thus, it can be concluded that the upper surface of She2p is required for the recognition of stem-loop containing RNAs as it was observed also for the basic helical hairpin motif (chapter 2.2.6).

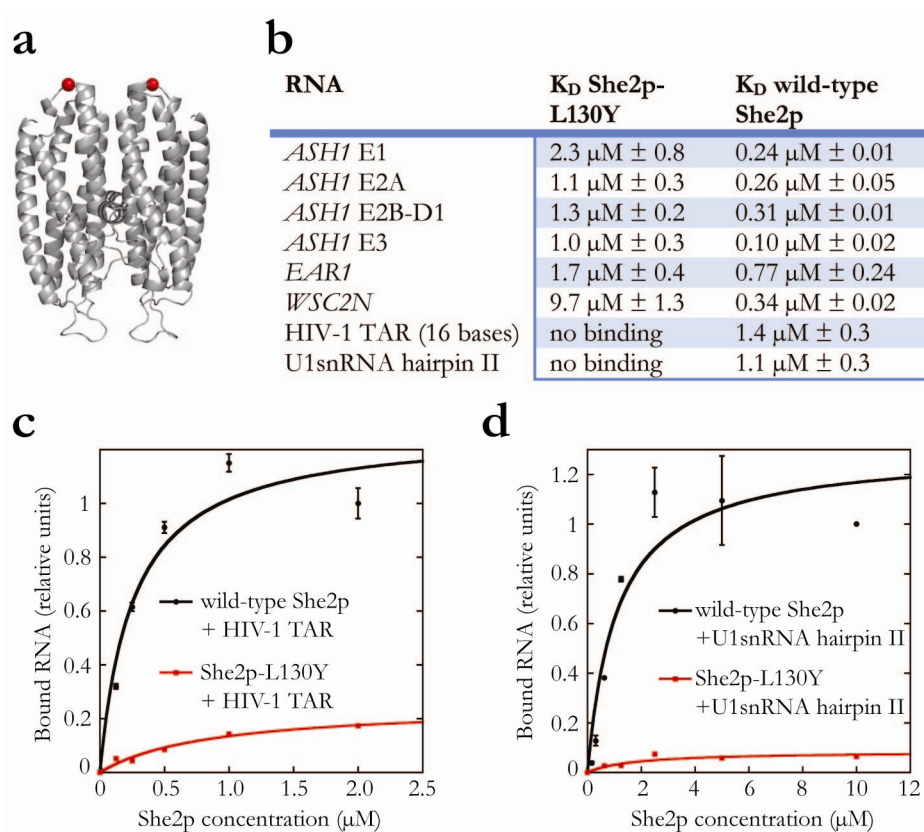


Figure 16: The upper surface region of She2p is required for recognition of stem-loop containing RNA. The cartoon in **a**) highlights leucine 130 (red) on both She2p monomers, which was mutated to tyrosine. The table in **b**) summarizes the RNA-binding experiments performed with She2p-L130Y in comparison to the respective binding constants obtained for wild-type She2p. She2p-L130Y binds significantly weaker to *ASH1* zipcodes and to *WSC2N* compared to wild-type She2p, but has an only two-fold reduced affinity to the *EAR1* zipcode. No significant binding was measured to the stem loops HIV-1 TAR and U1snRNA hairpin II up to a protein concentration of 32 μM. Graphs in **c**) and **d**) show the relative bound fraction by She2p-L130Y of HIV-1 TAR RNA and U1snRNA, respectively, compared to the relative bound RNA fraction by wild-type She2p.

2.2.8 Helix E is required for recognition of *ASH1* zipcodes

The mutant She2p- Δ helixE was created by deleting amino acids 174 to 183 in context of full-length She2p (Figure 14 c, g and Figure 17 a; see also chapter 2.1.2). To ensure that this deletion in She2p did not result in protein aggregation or instability, the size-exclusion chromatography profiles of wild-type She2p and She2p- Δ helixE were compared. As shown in Figure 17 b, wild-type and mutant protein eluted in a single peak from a Superdex 200 16/60 size-exclusion column at almost identical volumes of 58 ml and 57.8 ml, respectively, indicating that the mutant protein is properly folded. Filter-binding experiments revealed that deletion of helix E severely reduced She2p's affinity to all *ASH1*-zipcode elements and to the *WSC2N* zipcode by factor 4 to 20 (Figure 17 c). However, binding of She2p- Δ helixE to the *EAR1* zipcode was only modestly impaired (Figure 17 c). These findings suggest that the protruding helix E is important for the recognition of *ASH1* and *WSC2* mRNA-localization elements. Surprisingly, deletion of helix E had no significant influence on binding to the unrelated stem-loop containing RNAs HIV-1 TAR and U1snRNA (Figure 17 c). These results suggest that helix E in She2p specifically recognizes at least a subset of bud-localizing mRNAs and might function as a cargo sensor to discriminate between target and non-target RNAs. However, because *EAR1*-zipcode binding is only modestly impaired upon deletion of helix E, this sensing mechanism may not apply to all bud-localizing mRNAs in *S. cerevisiae*.

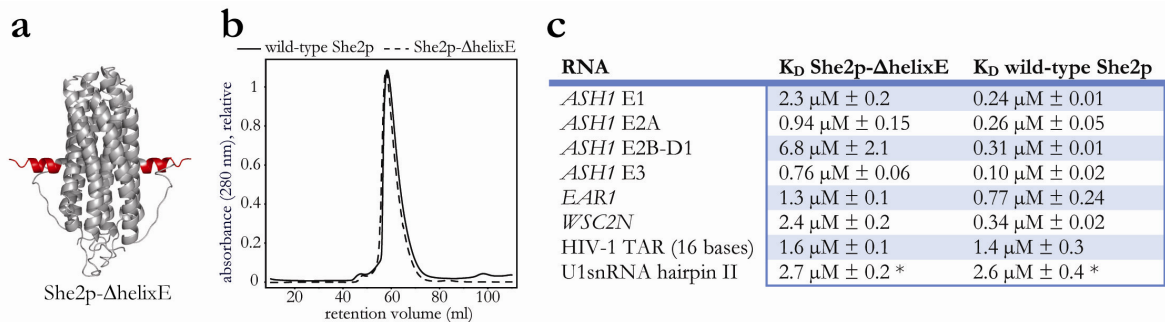


Figure 17: Helix E in She2p is required for *ASH1*-zipcode recognition. The cartoon in **a**) highlights the protruding helix E (red) on both She2p monomers, which was deleted to create She2p- Δ helixE. **b**) Comparative size-exclusion chromatography of wild-type She2p and She2p- Δ helixE shows that both variants elute at the same volume from a Superdex 200 16/60 column, suggesting intact folding of She2p- Δ helixE. The table in **c**) summarizes the RNA-binding experiments performed with She2p- Δ helixE in comparison to the respective binding constants obtained for wild-type She2p. Deletion of helix E results in a four- to 20-fold reduced affinity to *ASH1* and *WSC2N* zipcodes. The binding to the *EAR1* zipcode is only two-fold reduced, whereas the unrelated stem-loop RNAs HIV-1 TAR and U1snRNA hairpin II are equally bound by She2p- Δ helixE and wild-type She2p. Affinity to U1snRNA always varied between different sets of RNA-binding experiments with different She2-mutant proteins. For that reason, the results obtained from comparative experiments performed in parallel were included (marked by an asterisk).

2.2.9 The She2p C-terminus also contributes to *ASH1*-mRNA binding

In order to investigate the role of the She2p C-terminus in RNA binding, a variant was created lacking the C-terminal 6 amino acids of She2p ($\Delta 241$ -246) (Figure 14 d, h and Figure 18 a). Size-exclusion chromatography of this variant, termed She2p- Δ C-term, revealed no signs of aggregation or instability (Figure 18 b). When the elution profiles of wild-type She2p and She2p- Δ C-term were compared, a slight shift in the retention volume towards a smaller molecular weight was observed, which might reflect impaired oligomerization of She2p (Figure 18 b; see chapter 2.2.14). In filter-binding experiments, She2p- Δ C-term showed a significantly reduced affinity to the *ASH1*-E1, E2B and E3 zipcodes and binding to the *ASH1*-E2A zipcode could not be measured up to a protein concentration of 32 μ M (Figure 18 c). The binding of She2p- Δ C-term to the *EAR1* and *WSC2N* zipcodes remained unaffected or was even slightly increased for the unrelated HIV-1 TAR RNA and U1snRNA hairpin II (Figure 18 c). From these experiments it can be concluded that the flexible C-terminus of She2p contributes to *ASH1*-mRNA binding, but seems not to be involved in recognition of the bud-localizing mRNAs *EAR1* and *WSC2* or of unrelated stem-loop RNAs.

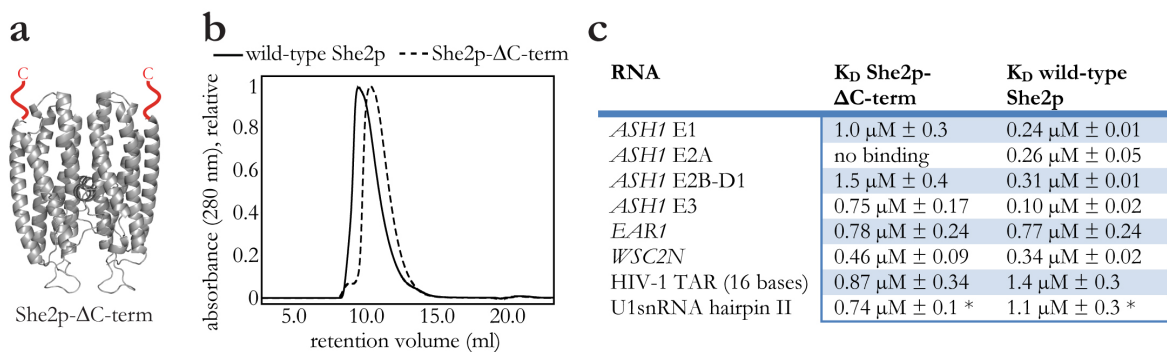


Figure 18: The C-terminus of She2p contributes to *ASH1*-zipcode binding. The cartoon in **a**) marks the position of the flexible C-terminus (red), which was not part of the She2p-crystal structure. **b**) Comparative size-exclusion chromatography of wild-type She2p and She2p- Δ C-term shows a slight shift in the elution volume of She2p- Δ C-term towards a smaller molecular weight. However, no aggregation of the mutant protein could be observed. In this experiment, a Superose 12 10/300 GL column was used. The table in **c**) summarizes the RNA-binding experiments performed with She2p- Δ helixE in comparison to the respective binding constants obtained for wild-type She2p. Deletion of the She2p C-terminus results in a significantly reduced affinity to *ASH1* zipcodes. Binding to the *EAR1* and the *WSC2N* zipcodes as well as binding to unrelated RNAs HIV-1 TAR and U1snRNA hairpin II remains almost unaffected by the C-terminal truncation. Affinities to U1snRNA varied between different sets of RNA-binding experiments with different She2-mutant proteins. The results obtained from comparative experiments performed in parallel are shown (marked by an asterisk).

RESULTS

In summary, using *in vitro* RNA-binding experiments four regions in She2p were identified to which two distinct RNA-binding properties of She2p could be attributed. General binding of RNA stem-loop structures is mediated by the basic helical hairpin motif (chapter 2.2.6). In addition, the upper surface region of She2p is required for the binding of RNAs that contain stem-loop secondary structures (chapter 2.2.7). In contrast, the protruding α -helix E and the very C-terminus of She2p mediate the recognition of the specific target mRNA *ASH1* (chapters 2.2.8 and 2.2.9). This implies that *ASH1* is bound by She2p in a different manner than unspecific non-target RNAs. However, She2p binding to the bud-localizing mRNAs *EAR1* and *WSC2* was not affected by the C-terminal deletion and only *WSC2N* binding was clearly dependent on helix E, suggesting that She2p might bind specific target RNAs in at least two distinguishable ways.

In the next step, the possible roles of the identified She2p-surface regions in cytoplasmic She2p and *ASH1*-mRNA localization, and on transport-complex assembly were analyzed *in vivo*.

2.2.10 Reduction of basic RNA binding abolishes bud-tip localization of She2p *in vivo*

Since She2p links more than 24 different zipcode-containing RNA cargoes to the motor protein Myo4p (Oeffinger et al., 2007; Shepard et al., 2003; Takizawa and Vale, 2000), She2p's ability to bind RNA-stem loops should be a prerequisite for the mRNA-localization process in *S. cerevisiae*. *In vitro*, selective mutation of the amino acids N36 and R63 in the basic helical hairpin motif completely abolished binding to specific and non-specific RNAs (chapter 2.2.6). Recently, it was shown that the double mutant She2p-N36S-R63K accumulates in the nucleolus of *S. cerevisiae* indicating that She2p is trapped in the nuclear subcompartment whenever it is unable to bind its target RNA (Du et al., 2008). Thus, cytoplasmic localization cannot be assessed for She2p-N36S-R63K.

The upper surface region of She2p was also shown to be required for basic RNA binding *in vitro* (chapter 2.2.7). Cytoplasmic localization of She2p-L130Y in comparison to the localization of wild-type She2p was analyzed by indirect immunofluorescence using a polyclonal α -She2p antibody. Yeast strains were generated that express either wild-type or mutant She2p under its endogenous promoter. Western-blot analysis of yeast-cell extracts confirmed an equal expression level of the respective proteins *in vivo* (Figure 19).

RESULTS

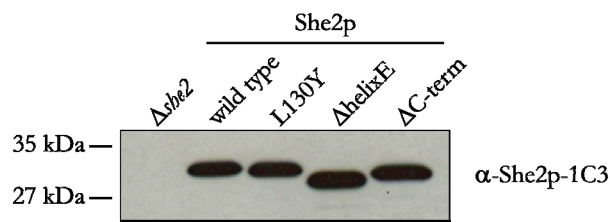


Figure 19: She2p mutants used for immunostaining experiments are expressed at wild-type She2p level *in vivo*. Cell extracts of wild-type She2p- and mutant She2p-expressing yeast strains were prepared and subjected to western-blot analysis using the monoclonal α -She2p-1C3 antibody. All three She2p mutants are expressed at levels comparable to wild-type She2p.

In immunostaining experiments I observed that wild-type She2p localizes in 69.2 % of all cells ($n = 3 \times \geq 100$ cells) to the tip of the daughter cell (Figure 20 a-c, m). In contrast, the RNA-binding mutant She2p-L130Y localized to the bud tip in only 2.9 % of inspected cells ($n > 100$ cells) (Figure 20 d-f, m). This finding indicates that disruption of basic RNA binding by mutating the upper surface of She2p results in mislocalization of She2p. This further suggests that also the localization of She2p-dependent mRNA should be impaired.

2.2.11 Selectively impaired *ASH1*-mRNA binding results in mislocalization of She2p *in vivo*

In vitro, binding of She2p to the target mRNA *ASH1* was shown to depend on the protruding α -helix E and the very C-terminal amino acids of She2p (chapters 2.2.8 and 2.2.9). However, She2p's affinity to the target mRNA *EAR1* was only moderately impaired by deletion of helix E and even not affected by the She2p C-terminal deletion. These findings let me assume that the *in vitro* observed specificity for target mRNA might be applicable only to *ASH1* mRNA or to a subset of *ASH1*-like mRNAs. *In vivo* transport of other bud-localizing mRNAs, like the *EAR1* mRNA, might be unaffected by these mutations. However, such a scenario would require proper She2p localization at the bud tip. Thus, a loss of She2p localization *in vivo* would be a good indication for a total loss of mRNA localization.

In order to analyze the general roles of helix E and the C-terminus on She2p localization, yeast strains were generated expressing She2p- Δ helixE or She2p- Δ C-term under the control of the endogenous promoter. Western-blot analysis of yeast-cell extracts confirmed that expression levels of the mutant proteins are comparable to the levels of wild-type She2p (Figure 19). Indirect immunofluorescence microscopy using an α -She2p antibody revealed that, in contrast to wild-type She2p, She2p- Δ helixE completely failed to localize to the bud tip (0 % localization in $n > 100$ cells) (Figure 20 g-i, m). Similarly, She2p- Δ C-term localized in

RESULTS

only 1 % of inspected cells ($n > 100$ cells) to the tip of the daughter cell (Figure 20 j-l, m). These experiments show that selective impairment of *ASH1*-mRNA binding results in a complete mislocalization of She2p *in vivo*. Since no RNA-binding mutant showed any significant bud-tip localization, it can be concluded that mRNA localization might be generally affected in these mutants.

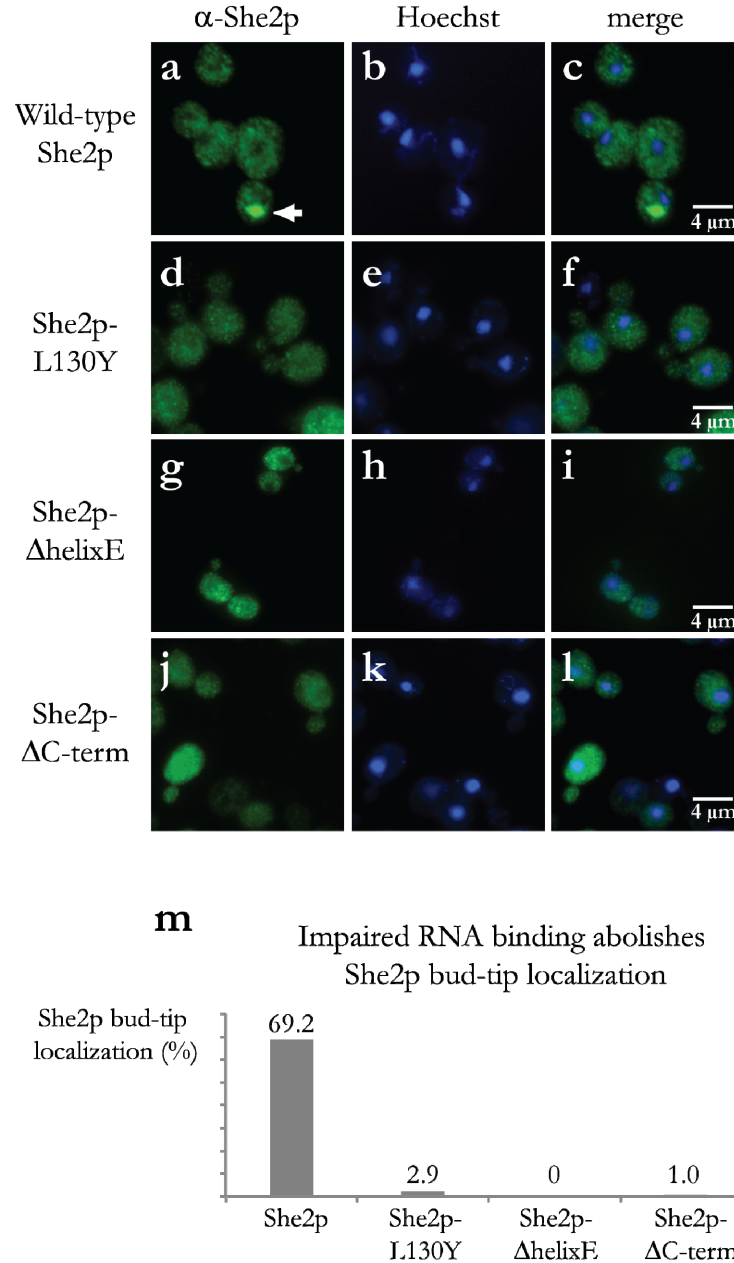


Figure 20: Impaired RNA binding by She2p results in disrupted bud-tip localization of She2p. Indirect immunofluorescence against She2p revealed that wild-type She2p efficiently localizes to the bud tip (a-c) in 69.2 % of inspected cells ($n = 3 \times \geq 100$ cells) (m). In contrast, localization of the RNA-binding deficient She2p variants She2p-L130Y (d-f), She2p-ΔhelixE (g-i), or She2p-ΔC-term (j-l) was never observed in more than 2.9 % of all budding cells ($n \geq 100$ cells) (m). She2p was immunostained with α-She2p (323/4) (a, d, g, j) and nuclei were visualized by Hoechst staining (b, e, h, k). Images in c, f, i, l display merged signals of the respective α-She2p and nuclei stainings. The chart in m represents the results of counting the number of cells with She2p bud-tip localization. For analysis, only budding yeast cells were counted. In *she2Δ* cells, no antibody staining was observed (not shown).

2.2.12 Impaired *ASH1*-mRNA binding by She2p results in mislocalization of *ASH1* mRNA *in vivo*

To further confirm the loss of *ASH1*-mRNA localization in She2p- Δ helixE mutant cells, *in situ* hybridization experiments using oligonucleotides directed against *ASH1* mRNA were performed. This collaborative experiment was carried out by Maria Schmid from the laboratory of Prof. Ralf-Peter Jansen. Since these are not my own results, the outcome of this experiment is described here, but the respective figure is shown in the appendix (Figure A4). The experiment revealed that, in contrast to a wild-type She2p background, *ASH1*-mRNA localization was completely abolished upon deletion of helix E (0 % localization in $n \geq 100$ cells; Figure A4). Diffuse distribution of *ASH1* mRNA is consistent with the observation that She2p- Δ helixE is uniformly distributed and not localized to the bud tip. Considering the findings, that none of the RNA-binding mutants localized to the bud tip (chapter 2.2.10) and that *ASH1* mRNA is delocalized in one of these mutants, no further *in situ* hybridization experiments against *ASH1* mRNA were performed.

2.2.13 Impaired RNA binding by She2p affects translocation-complex assembly

Immunostaining and *in situ*-hybridization experiments described in chapters 2.2.11 and 2.2.12 suggested that the identified surface regions in She2p have to act in an orchestrated fashion to allow for She2p localization and thus, for mRNA localization. However, the inability of She2p RNA-binding mutants to localize to the bud tip raised the question if these mutants still assemble with She3p and Myo4p into translocation complexes. To analyze cytoplasmic transport-complex assembly, myc-tagged She2p was subjected to immunoprecipitation with a monoclonal α -myc 9E11 antibody coupled to magnetic Protein-G beads. By western blotting, co-immunoprecipitation of the adapter protein She3p was analyzed and the respective She2p-associated fraction of She3p was quantified. The fractions that co-immunoprecipitated with She2p were not analyzed for the presence of the motor protein Myo4p, since She3p and Myo4p have been shown to form a stable cytoplasmic complex (Böhl et al., 2000; Heuck et al., 2007; Long et al., 2000; Takizawa and Vale, 2000). *she2 Δ* yeast cells were transformed with plasmids encoding for C-terminally myc-tagged She2p variants, which were expressed under its endogenous promoter. For visualization of She3p, it was genomically tagged with an N-terminal HA tag. Western blotting against She2p in yeast strains expressing the respective

RESULTS

She2p variants revealed that the expression level of the She2p-mutant proteins was comparable to the level of wild-type She2p (Figure 21).

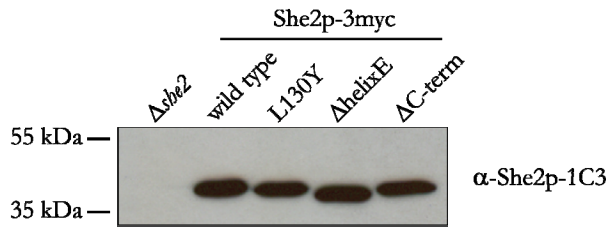


Figure 21: She2p mutants used for Co-IP experiments are expressed at wild-type She2p level *in vivo*. Cell extracts of wild-type She2p-3myc- and mutant She2p-3myc-expressing yeast strains were prepared and subjected to western-blot analysis using the monoclonal α -She2p-1C3 antibody. All three She2p mutants are expressed at levels comparable to wild-type She2p.

As shown in Figure 22, wild-type She2p efficiently co-precipitated She3p (Figure 22 a, upper panel, IP lanes). The different She2p mutants co-precipitated She3p less efficiently, which is indicated by a lower She3p-signal intensity in the respective IP fractions (Figure 22 a, upper panel, IP lanes). The lower panel in Figure 22 a shows the She2p input (Input lanes) and the unbound She2p fraction (Flow-through lanes). Apparently, unequal amounts of the respective She2p variants were bound to the beads, since the She2p signals in the flow-through fraction vary (Figure 22 a, lower panel, flow-through lanes). In order to draw quantitative conclusions of this experiment, the relative amount of She3p that is bound to She2p had to be calculated. Therefore, the She2p signals in the flow-through fractions were subtracted from the She2p-input signals (Figure 22 a, lower panel). The resulting value represented the total bead-bound fraction of She2p, which served as the effective input for co-immunoprecipitation. This value was then used to normalize She3p signals. The She3p amount that was co-precipitated by wild-type She2p was set to 100 % (Figure 22 b). Compared to the She3p association with wild-type She2p, She2p-L130Y assembled into 48.6 % of She3p-containing transport complexes (Figure 22 b). However, when considering the complete loss of She2p-L130Y bud-tip localization observed *in vivo* (Figure 20), it has to be concluded that the complexes assembled with She2p-L130Y are unstable and are not efficiently transported to the bud tip. In contrast, the She2p variants She2p- Δ helixE and She2p- Δ C-term, which are impaired in specific mRNA binding, failed to associate with She3p *in vivo* above background levels (Figure 22 b). This suggests that specific mRNA target recognition by She2p triggers the assembly of functional mRNA-translocation complexes.

Taken together, the co-immunoprecipitation experiments show that the She2p:She3p association is significantly impaired when She2p's affinity to RNA is reduced, indicating that mRNA binding by She2p is a prerequisite for mRNP assembly.

RESULTS

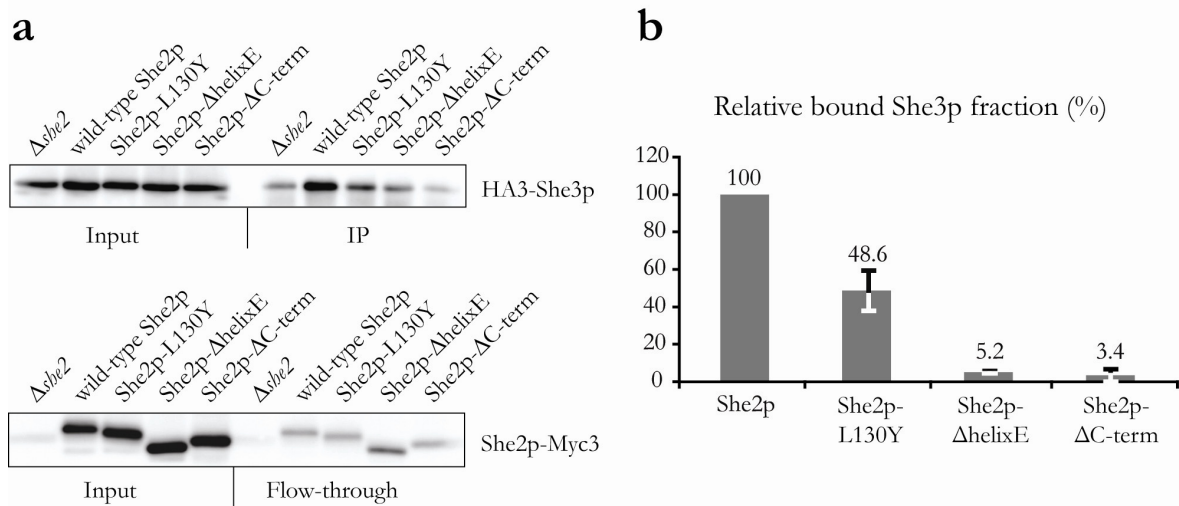


Figure 22: mRNP assembly is impaired in She2p RNA-binding mutants. **a)** Co-immunoprecipitation experiments of myc-tagged She2p and HA-tagged She3p. She2p and She3p were monitored by western blotting using α -myc 9E10 and α -HA 3F10 antibodies (lower and upper panel, respectively). Input fractions show that She3p expression was not influenced in She2p-mutant expressing yeast cells (upper panel). Upper panel, right: Wild-type She2p efficiently co-immunoprecipitated She3p. Complex assembly was consistently impaired in all She2p mutants as indicated by a lower She3p-signal intensity in the IP fractions. The band in the $\Delta she2$ lane indicates unspecific background binding of She3p to protein G-beads. Lower panel, right: The amount of the She2p input and the unbound She2p-3myc fraction is shown (Input and Flow-through lanes). For quantification of the She2p:She3p association the total bead-bound fraction of She2p was calculated by subtracting the unbound She2p fraction from the She2p-input fraction (lower panel). The respective bead-bound fractions of She2p were used to normalize She3p signals. **b)** Quantification of two independent experiments revealed that She2p-L130Y assembles into destabilized transport complexes, whereas She2p- Δ helixE and She2p- Δ C-term almost failed to associate with She3p.

2.2.14 She2p forms homo-oligomers through the upper surface region

During protein purification I observed that wild-type She2p elutes from size-exclusion columns at a volume far above the calculated volume of a She2p dimer (Figure 23). The molecular weight of a She2p dimer is 57.2 kDa, which would correspond to a retention volume of about 81 ml from a calibrated Superdex S200 16/60 size-exclusion column. However, for wild-type She2p typically a retention volume of 60 ml was observed correlating with a molecular weight of 320 kDa (Figure 23). This finding raised the question whether the She2p dimer observed in the crystal structure indeed represents the native oligomerization state of She2p. In addition, I found that the variant She2p-L130Y elutes at a volume corresponding to a significant lower molecular weight of about 110 kDa (Figure 23), indicating that mutation of the upper surface region of She2p disrupts the potential oligomeric state.

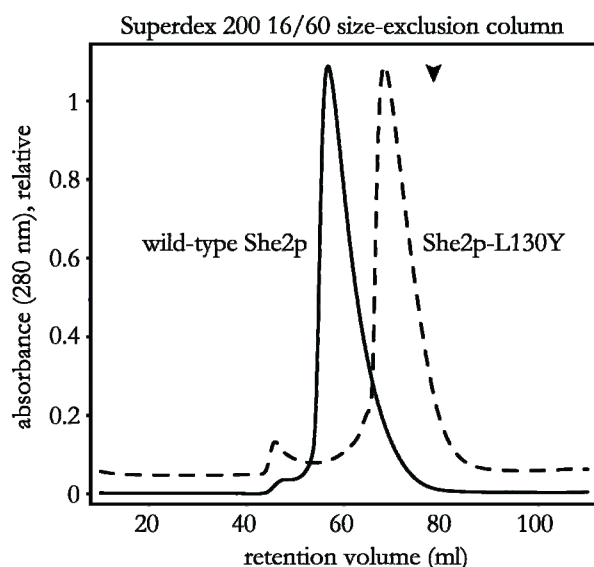


Figure 23: Preparative size-exclusion chromatography indicates a difference in molecular weight of wild-type She2p and She2p-L130Y. Wild-type She2p elutes from a Superdex 200 16/60 column at a volume of about 60 ml corresponding to a molecular weight of 320 kDa, indicating a higher oligomeric state than a She2p dimer (57.2 kDa). A significant shift in the elution volume of the She2p-L130Y variant towards a smaller molecular weight of 110 kDa was observed, suggesting an oligomerization defect. The theoretical retention volume of dimeric She2p (81 ml) is marked by an arrow head. The size-exclusion column was calibrated with a molecular weight marker kit.

2.2.14.1 Dynamic Light Scattering indicates large She2p particles

Using Dynamic Light Scattering (DLS), I determined the hydrodynamic radius (R_H) of wild-type She2p. The hydrodynamic radius represents the effective size of a molecule as detected by its diffusion. Thus, additional information on the She2p-particle size can be obtained. Wild-type She2p was measured at different concentrations in two different buffer compositions (HEPES and Tris based, respectively; Table 3). The measurements revealed that wild-type She2p indeed forms large particles with a R_H ranging from 9.5 nm to 13.5 nm (Table 3). It was observed that the particle size increases with higher protein concentrations, indicating a concentration-dependent oligomerization of She2p. Secondly, the overall size of the She2p oligomers seemed to significantly depend on the buffer composition, since smaller R_H values were obtained in a Tris-buffered solution. However, a concentration-dependent increase in particle size was also observed in Tris buffer. For comparison DLS measurements were performed with the variant She2p-L130Y, which was shown to elute at a smaller molecular weight from a size-exclusion column (Figure 23). She2p-L130Y was analyzed only in a Tris-based buffer. In contrast to wild-type She2p, She2p-L130Y forms significantly smaller particles with R_H values ranging from 5.3 nm to 7.4 nm (Table 3), indicating that the L130Y mutation indeed disrupts She2p-oligomer formation. However, She2p-L130Y was measured at a lower concentration range, but even at comparable protein concentrations a clearly different R_H was observed (Table 3).

RESULTS

Protein	Buffer	Conc. (mg/ml)	R _H (nm)
wild-type	10 mM HEPES,	3.5	11.5
She2p	125 mM NaCl, 1 mM DTT; pH 7.5	7.5	13.5
wild-type	20 mM Tris, 125 mM	3.9	9.5
She2p	NaCl, 1 mM DTT; pH 7.5	7.7	11.3
		14.4	12.4
She2p-L130Y	20 mM Tris, 125 mM	1.7	5.3
	NaCl, 1 mM DTT; pH 7.5	2.3	6.2
		3.1	7.4

Table 3: Dynamic light scattering indicates She2p-oligomer formation. DLS data show that wild-type She2p forms significantly larger particles than the mutant She2p-L130Y, suggesting that the upper surface region of She2p is involved in oligomerization.

Taken together, the DLS experiments suggest that She2p oligomerizes to higher molecular weight particles, which is dependent on the protein concentration and the surrounding solution. The oligomerization appears to be specific, because it can be impaired by mutating a single amino acid (leucine 130) in the upper surface region of She2p.

2.2.14.2 Analytical ultracentrifugation identifies She2p decamers that assemble through the upper surface region

In order to analyze the oligomerization state of She2p more precisely, I decided to use the technique of analytical ultracentrifugation (AUC). These experiments were performed in collaboration with Dr. Klaus Richter from the Department of Chemistry at the Technical University Garching, who also analyzed the data. Since the AUC experiments proved to be essential for subsequent experiments, the results are shown in this section.

To characterize the oligomeric state of both wild-type She2p and She2p-L130Y, analytical ultracentrifugation experiments were performed. Both She2p variants were pre-treated with RNase A to prevent unspecific RNA-mediated oligomerization and then analyzed in the Tris-based size-exclusion buffer 2 (see chapter 4.9.4). Sedimentation-velocity experiments performed with wild-type She2p at a concentration of 16 μ M revealed one single oligomeric species (Figure 24 a). From the sedimentation profile of wild-type She2p (appendix, Figure A5) a sedimentation coefficient of 9.6 Svedberg (S) was determined (Figure 24 a). Subsequently, sedimentation-velocity experiments were performed with She2p-L130Y, for which a smaller particle size and thus, a smaller sedimentation coefficient were assumed. She2p-L130Y was analyzed in three different concentrations (4.9 μ M, 12.9 μ M, 32.0 μ M) to identify potentially different oligomeric species. Indeed, the mutant protein sedimented in a

RESULTS

concentration-dependent manner between 4.2 S and 7.5 S (Figure 24 b) with an average of 5.6 S (Figure 24 a). However, even at the highest protein concentration the sedimentation coefficient of She2p-L130Y was still substantially smaller than the one obtained for wild-type She2p at an intermediate concentration of 16 μ M (Figure 24 a, b). These results clearly indicate different oligomeric states of wild-type and mutant She2 protein.

In order to calculate the molecular weight of wild-type and mutant She2p, sedimentation-equilibrium experiments were performed. Both proteins were measured at nine different concentrations ranging from 1.7 μ M to 30.7 μ M (wild-type She2p) and 1.9 μ M to 32.0 μ M (She2p-L130Y) (see chapter 4.10.3) to investigate concentration-dependent oligomerization events. For wild-type She2p, a molecular weight of approximately 290 kDa was determined (Figure 24 c), which is consistent with the molecular weight derived from size-exclusion chromatography and also with the large hydrodynamic radius measured by DLS. In contrast to the DLS measurements, no concentration-dependent oligomerization could be observed in the sedimentation-equilibrium experiments. However, the protein concentration range used in both experiments significantly differed. The maximum concentration measured in AUC experiments was 0.9 mg/ml, whereas due to the low detection limit in DLS experiments higher protein concentrations had to be used. This might suggest that at a certain protein concentration She2p tends to form non-specific higher molecular-weight complexes. Considering a molecular weight of a She2p dimer of 57.2 kDa the oligomer observed in AUC experiments corresponds to a complex consisting of ten She2p monomers. Since it is well-proved that She2p forms dimers with a buried surface area of 2015 \AA^2 per monomer (Niessing et al., 2004), a complex built of five She2p dimers can be suggested.

Sedimentation-equilibrium experiments performed with She2p-L130Y revealed that it oligomerizes concentration-dependently as it was already observed in the DLS experiment and in the sedimentation-velocity experiment. For an intermediate protein concentration of 11.1 μ M a molecular weight of 76.1 kDa was calculated (Figure 24 d), which corresponds to an oligomerization state between a dimer (57.2 kDa) and a tetramer (114.4 kDa). From this finding it can be concluded that mutation of leucine 130 interferes with She2p oligomerization.

RESULTS

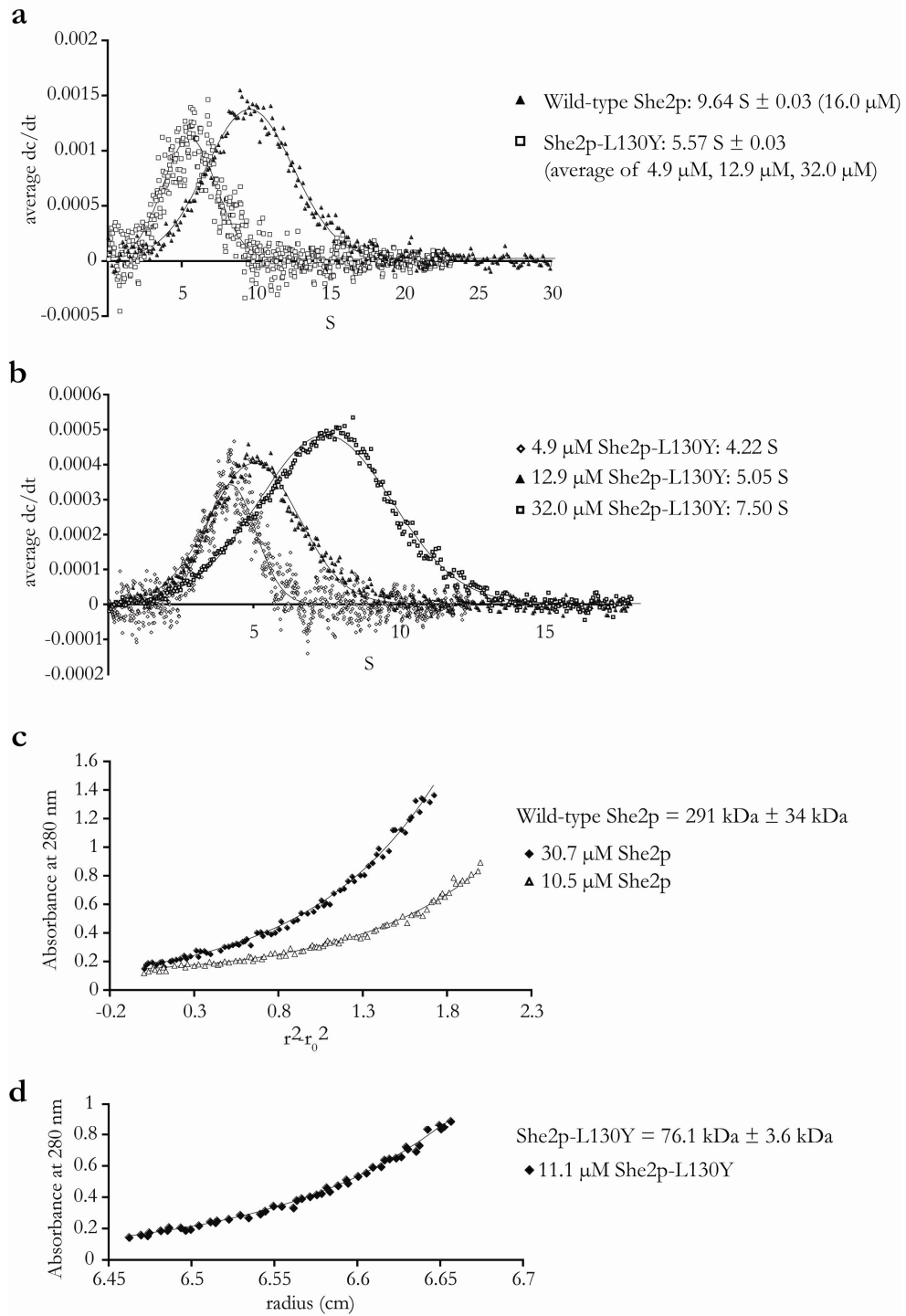


Figure 24: Analytical-ultracentrifugation experiments suggest that She2p oligomerizes via its upper surface region. **a)** In sedimentation-velocity experiments wild-type She2p sediments at $9.6 S$. In contrast, She2p-L130Y sediments concentration-dependently with an average of $5.6 S$. The reduced sedimentation coefficient for She2p-L130Y is consistent with a reduced molecular weight. **b)** Sedimentation-velocity experiments with She2p-L130Y at different protein concentrations show a concentration-dependent change in sedimentation. However, even at a protein concentration of $32 \mu M$, the sedimentation coefficient of wild-type She2p is not reached. **c)** Sedimentation-equilibrium experiments with wild-type She2p at different protein concentrations indicate a concentration-independent oligomer of 290 kDa . Shown are two representative curves that could be fitted to the same molecular weight. **d)** Sedimentation-equilibrium experiments with She2p-L130Y at an intermediate concentration of $11.1 \mu M$ revealed a molecular weight of 76 kDa , confirming that higher molecular weight oligomers are disrupted by this mutation.

RESULTS

In summary, analytical ultracentrifugation experiments confirmed that She2p forms higher-order oligomers via the upper surface region of the She2p dimer. Together with the RNA-binding experiments and *in vivo* studies, these results indicate that She2p oligomerization is important for the binding of stem-loop containing RNAs and that, in turn, RNA binding by She2p oligomers is required for efficient mRNP translocation to the bud tip.

2.2.14.3 Electron microscopy suggests She2p oligomers smaller than a decamer

If the She2p oligomer of 290 kDa observed in analytical ultracentrifugation experiments was true, visualization of the complex by electron microscopy should be possible. Therefore, negative-stain images of wild-type She2p were recorded by the laboratory of Prof. Roland Beckmann (Gene Center Munich). Unfortunately, uniform particles could not be identified, but some of the observed particles showed a stretched shape of about ten nanometers in length (not shown). However, for a protein of 290 kDa a much larger particle size would be expected. The protein concentrations that are used in electron-microscopy experiments are very low (100 µg/ml, corresponding to 3.5 µM She2p). This might be an indication that She2p forms smaller particles at more physiological concentrations, which are below the detection limit of the biophysical methods that were used so far.

2.2.15 She2p is tetrameric at physiological protein concentrations

In order to analyze She2p oligomerization at near-physiological protein concentrations I performed analytical size-exclusion experiments with wild-type She2p and She2p-L130Y. For the analysis, a Superose 12 10/300 size-exclusion column was used, which was calibrated to allow for molecular-weight determination. The estimated physiological concentration of She2p *in vivo* was calculated to be 230 nM (Niessing et al., 2004). Therefore, various protein dilutions were chosen covering concentrations below and above the assumed physiological concentration (Figure 25). The proteins were purified under high-salt conditions to prevent any unspecific RNA-mediated oligomerization. The respective protein samples were applied to the column and analyzed in size-exclusion buffer 3 (chapter 4.9.4). To account for sample dilution on the size-exclusion column, She2p-containing peak fractions were integrated over the peak volume. Thereby, the approximate protein concentration on the column could be

RESULTS

calculated. As shown in Figure 25 a and b, the respective applied and estimated protein concentrations differed roughly by factor 2.5 to 3.

Considering sample dilution, wild-type She2p formed a distinct oligomeric complex up to an experimental concentration of 0.78 μM with an average molecular weight of 125.5 kDa corresponding to a She2p tetramer ($4 \times 28.6 \text{ kDa} = 114.4 \text{ kDa}$) (Figure 25 a, c). Importantly, this tetrameric complex was determined for protein concentrations around or even above the physiological She2p concentration of 230 nM. This finding suggests that She2p forms larger oligomers than the dimer observed in the crystal structure and that a She2p tetramer might be the oligomerization state *in vivo*. Up to an experimental protein concentration of 3.5 μM , the molecular weight of wild-type She2p slightly increased to 145 kDa (Figure 25 a, c). However, at higher concentrations She2p formed larger complexes of almost 200 kDa (Figure 25 a, c), which reflects the concentration-dependent oligomerization observed in dynamic light scattering (chapter 2.2.14.1).

In contrast, She2p-L130Y failed to form oligomers corresponding to a putative tetramer (Figure 25 b, c). Instead, up to an experimental concentration of 330 nM a molecular weight of 58 kDa was calculated (Figure 25 b, c), which exactly matches the molecular weight of a She2p dimer (57.2 kDa). This result strongly suggests that mutation of leucine 130 at the upper surface region disrupts She2p-tetramer formation at physiological protein concentrations. With higher protein concentrations, an increase of molecular weight of She2p-L130Y was observed, being consistent with previous observations (chapters 2.2.14.1 and 2.2.14.2). But not even at the highest protein concentration measured the oligomeric state of wild-type She2p was reached (Figure 25 c).

In previous chapters, a number of functional defects were observed for She2p-L130Y. This mutant has an impaired RNA binding (chapter 2.2.7), it does not localize to the bud tip (chapter 2.2.10), and it shows defects in transport-complex assembly *in vivo* (chapter 2.2.13). Further I have shown by various techniques that mutation of leucine 130 to tyrosine disrupts She2p oligomerization. Therefore, I conclude that wild-type She2p forms tetramers at physiological concentrations via its upper uncharged surface region and that these tetramers are functionally relevant *in vivo*.

RESULTS

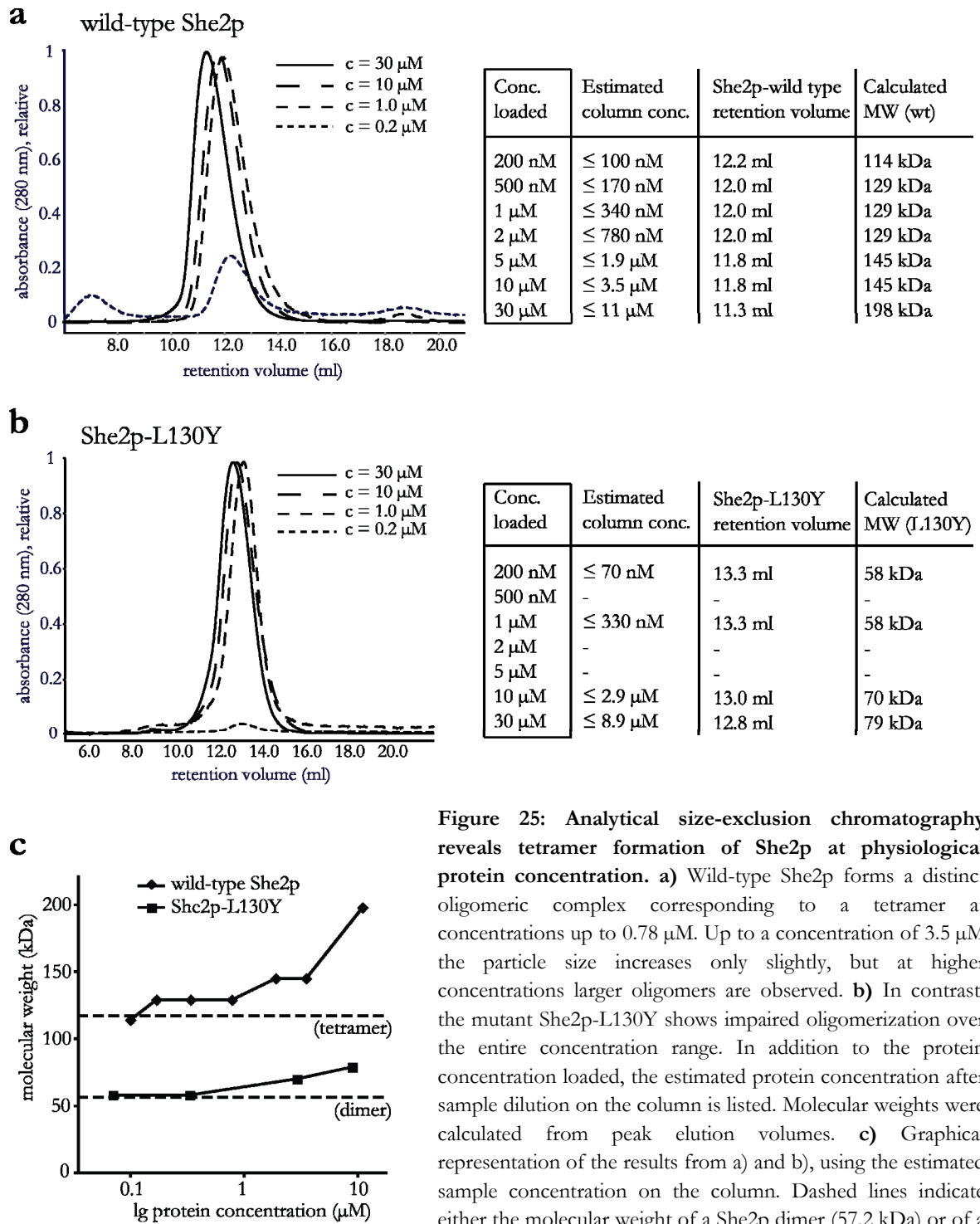


Figure 25: Analytical size-exclusion chromatography reveals tetramer formation of She2p at physiological protein concentration. **a)** Wild-type She2p forms a distinct oligomeric complex corresponding to a tetramer at concentrations up to 0.78 μM. Up to a concentration of 3.5 μM the particle size increases only slightly, but at higher concentrations larger oligomers are observed. **b)** In contrast, the mutant She2p-L130Y shows impaired oligomerization over the entire concentration range. In addition to the protein concentration loaded, the estimated protein concentration after sample dilution on the column is listed. Molecular weights were calculated from peak elution volumes. **c)** Graphical representation of the results from a) and b), using the estimated sample concentration on the column. Dashed lines indicate either the molecular weight of a She2p dimer (57.2 kDa) or of a tetramer (114.4 kDa).

2.2.16 Small Angle X-ray Scattering (SAXS) studies identify She2p tetramers

By using several techniques described above (dynamic light scattering, analytical ultracentrifugation, and analytical size-exclusion chromatography) I observed that She2p forms larger oligomers than the dimer described in the crystal structure (Niessing et al., 2004). In order to analyze the oligomeric state from a structural point of view, I performed Small Angle X-ray Scattering (SAXS) experiments with wild-type She2p and the oligomerization-defective variant She2p-L130Y.

SAXS is a powerful technique that can be used to obtain low-resolution structural information on biological macromolecules (reviewed in Putnam et al., 2007). SAXS experiments are performed in solution and therefore, the sample can be measured in a number of different buffer conditions, preferably in a buffer reflecting the physiological composition. In advantage to X-ray crystallography, SAXS can also provide structural information on flexible parts of a protein. Combining SAXS-scattering data with the crystal structure of a given protein provides a potential tool to obtain more accurate and complete models of protein complexes in solution. In addition, SAXS has been successfully used to determine the molecular masses of proteins with an average error below 10 % (Mylonas and Svergun, 2007). Scattering data of monodisperse samples are usually collected at the highest possible protein concentration without observing any aggregation of the protein. In order to obtain reliable information from the scattering data, the minimal protein concentration must be higher than 1 mg/ml, since otherwise the signal-to-noise ratio is too low.

2.2.16.1 Identification of suitable buffer conditions for SAXS measurements

At a protein concentration of 1 mg/ml (34.9 μ M) She2p oligomerizes into a larger complex than the tetramer that is formed at a physiological protein concentration (chapter 2.2.15). Therefore, a buffer condition had to be identified, in which She2p is present as a stable, physiological oligomer at protein concentrations higher than 1 mg/ml. Using DLS, a number of different buffers were screened for their influence on She2p's solubility and oligomerization (Table 4). Besides variations in salt or buffer composition, also several detergents were tested, which are known to function in SAXS experiments at concentrations below the critical micelle-forming concentration (<http://bl1231.als.lbl.gov/>). As shown in table 4, the individual buffer composition had a strong effect on She2p's particle size and its solubility. In most

RESULTS

conditions, a hydrodynamic radius (R_H) between 9 nm and 11 nm was observed (Table 4), which is comparable to the radii measured for wild-type She2p in previous DSL experiments (chapter 2.2.14.1). Aggregation of She2p was especially provoked by the addition of detergents to the solution. In contrast, in buffers with a high ionic strength (#2 and #4) She2p formed smaller particles with an R_H between 5.6 nm and 6.1 nm (Table 4). This finding indicates that She2p exists in different oligomeric states in solution and that She2p oligomerization strongly depends on the buffer condition. Importantly, the size of the She2p oligomer observed in buffers #2 and #4 increased only slightly with higher protein concentrations rendering these conditions promising for following SAXS measurements.

#	Buffer	Conc. (mg/ml)	R_H (nm)	Sample aggregation?	Suitable for SAXS?
1	20 mM Tris, 150 mM NaCl, 2 mM DTT; pH 7.5	2.6	8.7	-	+
		6.3	10.7	-	
2	20 mM Tris, 300 mM NaCl, 2 mM DTT; pH 7.5	2.6	5.7	-	+
		5.7	6.1	-	
3	20 mM Tris, 150 mM NaCl, 2 mM DTT, 5 % glycerol; pH 7.5	2.6	10.1	+	-
		4.9	12.0	-	
4	50 mM $\text{NaH}_2\text{PO}_4/\text{Na}_2\text{HPO}_4$, 150 mM NaCl, 2 mM DTT; pH 7.4	2.6	5.6	-	+
		6.5	5.9	-	
5	20 mM Tris, 150 mM NaCl, 2 mM DTT, 5 mM β -octyl glucoside; pH 7.5	2.4	10.0	-	-/+
		3.9	9.7	-/+	
6	20 mM Tris, 150 mM NaCl, 2 mM DTT, 0.1 % CHAPS; pH 7.5	2.1	10.0	-	+
		3.7	10.5	-	
7	20 mM Tris, 150 mM NaCl, 2 mM DTT, 1.5 mM Decyl- β -D-maltoside; pH 7.5	2.8	10.2	+	-
		6.2	11.2	+	
8	20 mM Tris, 150 mM NaCl, 2 mM DTT, 0.2 % DDMAB; pH 7.5	2.0	>63.7	+	-
		2.9	>23.1	+	
9	20 mM Tris, 150 mM NaCl, 2 mM DTT, 0.4 % DDAO; pH 7.5	2.2	>22.5	+	-
		3.0	>58.4	+	

Table 4: Dynamic light scattering with wild-type She2p identified suitable buffer conditions for SAXS experiments. Summarized are the results from DLS experiments that were performed with wild-type She2p at the indicated protein concentrations in different buffer conditions. Generally, two different particle sizes of average 10 nm (#1, #3, #5, #6, #7) and 6 nm (#2, #4) were detected, indicating that She2p oligomerization strongly depends on the surrounding solution. Addition of detergents often induced aggregation of She2p. In buffers with a higher ionic strength (#2, #4) significantly smaller particles were observed with a constant R_H at high protein concentrations.

2.2.16.2 SAXS confirms tetramer formation in solution

SAXS experiments were performed with wild-type She2p in five different conditions and with She2p-L130Y in three different conditions, respectively (Table 5). The scattering data were primarily analyzed with PRIMUS from the ATSAS program package for the presence of aggregation or intermolecular forces between the molecules. With exception of data collected in buffer #4, all scattering data were inappropriate for further processing due to concentration-dependent oligomerization or protein aggregation (Table 5). This finding demonstrates the sensitivity of SAXS measurements compared to the technique of dynamic light scattering, as in the DLS experiments no sample aggregation could be detected in these buffer conditions (chapter 2.2.16.1).

Wild-type She2p		She2p-L130Y		#	Protein buffer
SAXS data collection?	Data processing?	SAXS data collection?	Data processing?		
+	Not possible. Tends to aggregate.	+	Partially possible.	1	20 mM Tris, 150 mM NaCl, 2 mM DTT; pH 7.5
+	Partially possible.	+	Partially possible.	2	20 mM Tris, 300 mM NaCl, 2 mM DTT; pH 7.5
+	Yes	+	Yes	4	50 mM NaH ₂ PO ₄ /Na ₂ HPO ₄ , 150 mM NaCl, 2 mM DTT; pH 7.4
+	Not possible. Protein aggregation.	-	-	5	20 mM Tris, 150 mM NaCl, 2 mM DTT, 5 mM β -octyl glucoside; pH 7.5
+	Not possible. Protein aggregation.	-	-	6	20 mM Tris, 150 mM NaCl, 2 mM DTT, 0.1 % CHAPS; pH 7.5

Table 5: Summary of SAXS-scattering data collected for wild-type She2p and She2p-L130Y. The table summarizes in which buffer conditions SAXS scattering data were collected (“+”) and if the scattering data could be processed. “Partially possible” indicates that not all measured data gave a meaningful result and therefore, were not further processed. Numbering of the protein buffers is identical to Table 4. Please note, that for comparison of the SAXS data, only scattering data were completely processed that were obtained for both She2p variants in identical buffer conditions (buffer #4).

Analysis of the scattering curves of wild-type She2p and She2p-L130Y which were measured in sodium phosphate-containing buffer (#4), revealed no signs of aggregation or intermolecular forces between the particles (Figure 26). Numerous consecutive measurements of each sample revealed no sensitivity of She2p towards long exposure to high energy X-ray radiation. The scattering curves obtained at the different concentrations show a good overlap

RESULTS

over long distances. However, the respective scattering curves measured at high protein concentration differed to some extent from the data obtained at lower concentrations, which might indicate a concentration-dependent behavior of the proteins. Therefore, all scattering curves were analyzed individually.

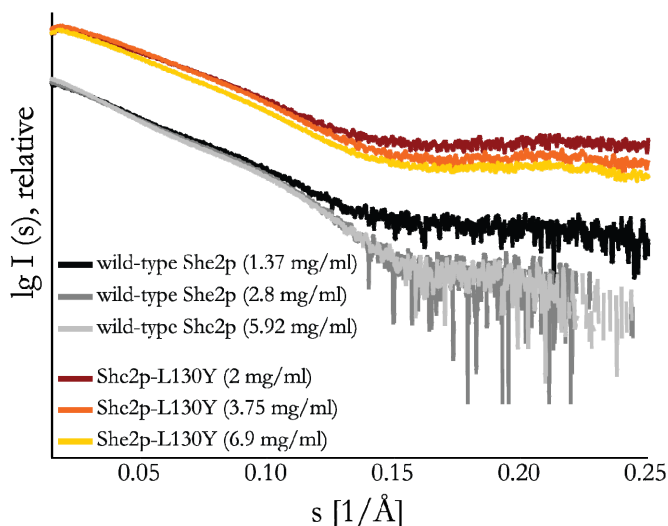


Figure 26: SAXS-scattering curves of wild-type She2p and She2p-L130Y measured in sodium-phosphate buffer. Shown are the scattering curves of both wild-type She2p and She2p-L130Y that were obtained for three concentrations each in sodium phosphate-containing buffer. The buffer was already subtracted from the individual curves. The scattering intensity I is plotted logarithmically against the scattering angle s .

The Kratky plot can be directly calculated from the scattering curves ($s^2 \cdot I(s)$ as a function of s) and allows for evaluating the folding of the measured samples. Typically, folded peptides show a peak with a roughly parabola-like shape. In contrast, unfolded or unstructured peptides lack this characteristic peak and are linear with respect to s at large angles or the curve is slightly increasing in the larger s -range. Analysis of wild-type She2p revealed a clear peak and no increase of the curve at large angles, indicating proper folding of the molecule (Figure 27 a). The Kratky plot of the mutant She2p-L130Y also shows a parabola-like peak, but in contrast to wild-type She2p, the protein might contain unstructured portions as the slope increases in the larger s -region (Figure 27 b).

RESULTS

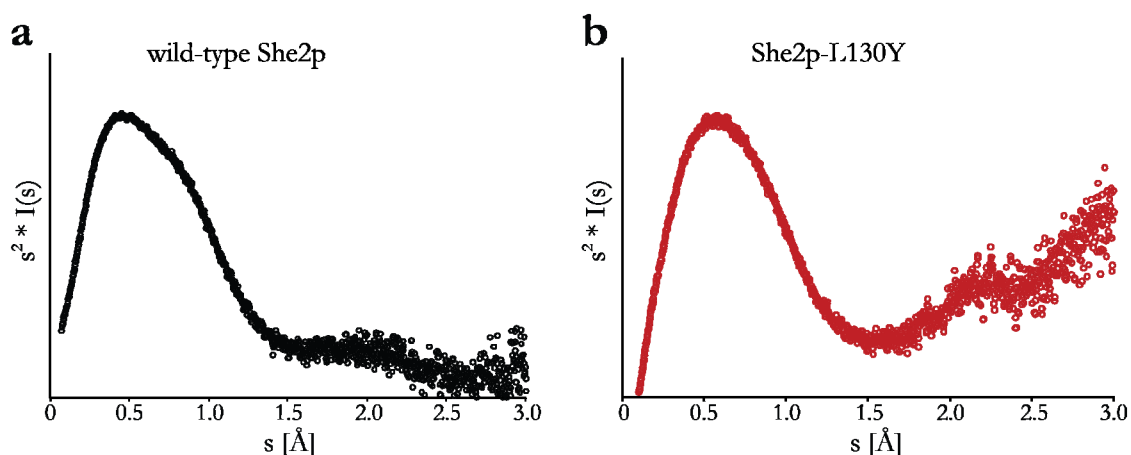


Figure 27: Kratky plot indicates proper folding of wild-type She2p and potentially unstructured regions in She2p-L130Y. The Kratky plot in **a)** shows a single peak, indicating proper folding of wild-type She2p. The plot in **b)** reveals that the mutant She2p-L130Y is predominantly folded, but might contain unstructured regions. Kratky plots were calculated directly from the scattering curves measured at intermediate protein concentrations (wild-type She2p: 2.8 mg/ml; She2p-L130Y: 3.75 mg/ml). Plotted is $s^2 \cdot I(s)$ as a function of the scattering angle s .

For molecular-mass determination of the individual complexes, the scattering intensities (I) were extrapolated to zero angle (I_0). The forward scattering intensity I_0 depends only on the square of the number of electrons and is independent on its shape. In an ideal sample the detected scattering is linearly related to the number of particles. With known sample concentration and by using reference samples like bovine serum albumin and lysozyme, the molecular weight of a molecule can be determined. For wild-type She2p a molecular weight between 107.9 kDa and 116.3 kDa was calculated (Table 6). This clearly corresponds to a tetramer (114.4 kDa) and confirms the results from analytical size-exclusion experiments (chapter 2.2.15). Corresponding analysis of She2p-L130Y showed that it associates concentration-dependently to complexes between 79.6 kDa and 102.2 kDa (Table 6), confirming the oligomerization defect of this mutant. Even at the highest protein concentration (6.9 mg/ml) She2p-L130Y fails to assemble into wild-type like oligomers, but instead forms an equilibrium between a dimeric and a tetrameric state (Table 6).

RESULTS

Protein	Conc. (mg/ml)	Conc. (μ M)	MW (kDa)	R_g (Å)	D_{max} (Å)	Average R_g (Å)	Average D_{max} (Å)
wild-type She2p	1.37	47.8	107.9	45.6	160.0	49.4 \pm 8.8	172.5 \pm 10.9
	2.8	97.7	115.8	49.5	180.0		
	5.92	206.5	116.3	53.2	177.5		
She2p-L130Y	2.0	69.8	79.6	36.8	137.5	38.8 \pm 2.2	138.6 \pm 5.9
	3.75	130.8	91.7	38.3	133.3		
	6.9	240.7	102.2	41.2	145.0		

Table 6: Structural parameters calculated from SAXS-scattering curves of wild-type She2p and She2p-L130Y. Summarized are the structural parameters that were derived from SAXS analysis of both wild-type and mutant She2p at indicated protein concentrations. R_g : radius of gyration; D_{max} : maximum dimension of the particle. By comparing the individual parameters of wild-type and mutant She2p, a significantly different overall size of both molecules can be observed.

In order to obtain information on the shape and the overall dimensions of She2p, the radii of gyration (R_g) were determined from the scattering curves. The radius of gyration is calculated from the Guinier approximation that is only valid for scattering at small angles ($s \rightarrow 0$). The Guinier plot, in which $\ln(I(s))$ is plotted versus s^2 , should give a linear function, from which I_0 and R_g can be extracted. The Guinier approximation is valid in the range of $s * R_g < 1.3$ and must be iteratively estimated. R_g is defined as the square root of the average squared distance of each scattering molecule from the center of the particle and therefore depends on the size and the shape of a particle. For wild-type She2p, an R_g ranging between 45.6 Å and 53.2 Å with an average of 49.4 Å was calculated (Table 6). In contrast, the radius of gyration obtained for the She2p-L130Y mutant was significantly smaller and was determined to be between 36.8 Å and 41.2 Å with an average of 38.8 Å (Table 6). Both proteins exhibited a concentration-dependent increase in R_g values, as it was already suggested in primary data analysis (see above). However, the respective average R_g values clearly differed from each other (Table 6), indicating different shapes of wild-type She2p and She2p-L130Y.

2.2.17 SAXS indicates that She2p forms an elongated tetramer

Calculation of the pair-distribution function $P(r)$ gives additional information on the overall shape of a molecule. The $P(r)$ function is calculated through a Fourier transform of the scattering curve $I(s)$ and provides information about the distances between electrons in the scattering particles. Importantly, the $P(r)$ function is valid over the entire s -range and thus, represents a real-space approximation of the scattering particles. Theoretically, the $P(r)$

RESULTS

function is zero at $r = 0$ and at $r \geq D_{\max}$. D_{\max} corresponds to the maximum linear dimension in the scattering particle. For transformation of the scattering data into the $P(r)$ function, the D_{\max} value has to be iteratively determined and the resulting $P(r)$ functions have to be analyzed for their fit to the experimental scattering. Subsequently, I determined the D_{\max} values of wild-type She2p to range between 160 Å and 180 Å, whereas for She2p-L130Y D_{\max} values between 133.3 Å and 145 Å were determined (Table 6). Consistently with the R_g values, also the D_{\max} values of both proteins increased concentration-dependently.

I calculated the $P(r)$ functions from all scattering data assuming the respective iteratively determined D_{\max} values (Table 6, Figure 28). The $P(r)$ function of wild-type She2p shows the typical skewed shape of an elongated molecule having a peak at short distances and an extended tail (Figure 28 a). In contrast, the $P(r)$ function of She2p-L130Y is rather symmetric with a single peak and shows the shape of a more globular molecule (Figure 28 a). It was observed that the $P(r)$ functions calculated for different concentrations of the respective proteins are not completely identical (Figure 28 b). Considering the differences in the D_{\max} values, on which the calculation of the $P(r)$ function relies, this finding was expected. Interestingly, the fact that the molecular weight of She2p-L130Y increases concentration-dependently without reaching the particle size of wild-type She2p can also be demonstrated with the $P(r)$ functions of She2p-L130Y. With increasing concentrations of She2p-L130Y, its $P(r)$ functions approached the shape of the wild-type She2p curves without ever merging with it (Figure 28 b).

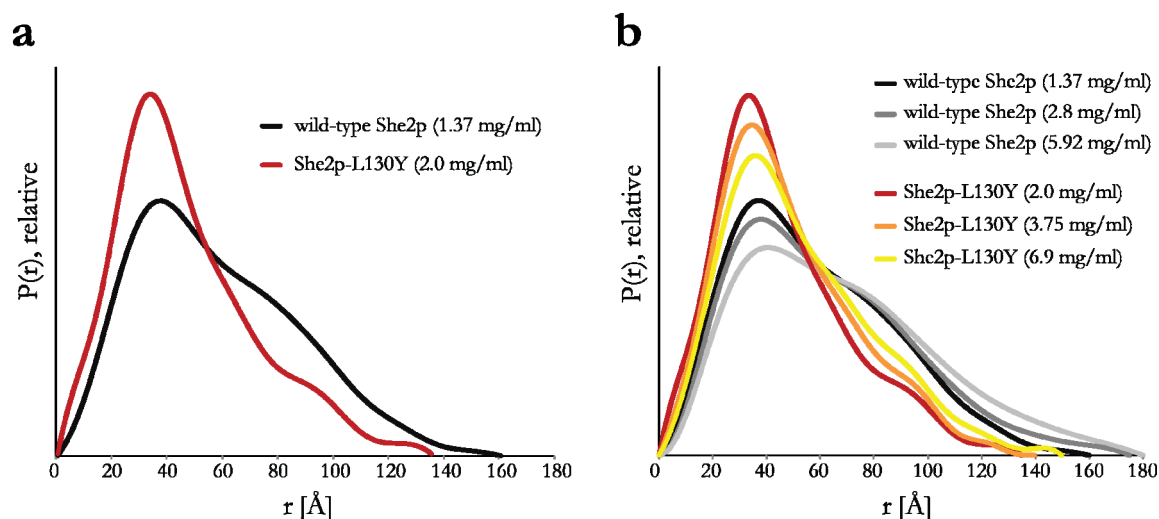


Figure 28: Pair-distribution functions of wild-type She2p and She2p-L130Y indicate different particle shapes. a) and **b)** show the $P(r)$ functions calculated from the SAXS-scattering data measured at the indicated protein concentrations of wild-type She2p and She2p-L130Y. **a)** The $P(r)$ function of the wild-type protein shows the typical shape of an elongated molecule, whereas the $P(r)$ function of She2p-L130Y suggests a globular molecule. **b)** $P(r)$ functions of wild-type and mutant protein demonstrate the concentration-dependent behavior of She2p-L130Y. With increasing concentration the $P(r)$ function approaches the shape of the wild-type $P(r)$ function, but does not merge with it. Also wild-type She2p shows a slight concentration dependency, but to a lower extent.

The experiments performed so far suggest that She2p forms tetramers, most likely comprising two dimeric subunits, and that these complexes can be disrupted by mutation of leucine 130. This amino acid is located in a conserved region on the upper hydrophobic surface of She2p. Therefore, She2p tetramers might form through the upper surface. In the published crystal structure contacts between She2p dimers could be observed (Figure 29) (Niessing et al., 2004). I decided to re-examine the potential role of these contacts for She2p-tetramer formation. Among the several crystallographic contacts, only two dimer interactions buried surface areas large enough to be potentially relevant in solution. One contact involves the basic helical hairpin-containing surface at the front of She2p and buries a surface area of 778 \AA^2 per dimer (Figure 29 a). This potential assembly was termed globular tetramer. The second dimer interaction involves the upper, uncharged surface region of She2p. It buries a surface area of 743 \AA^2 per dimer and was termed elongated tetramer (Figure 29 a, b). Leucine 130 is located exactly in this region and therefore, formation of an elongated tetramer was suggested (Figure 29 c, d).

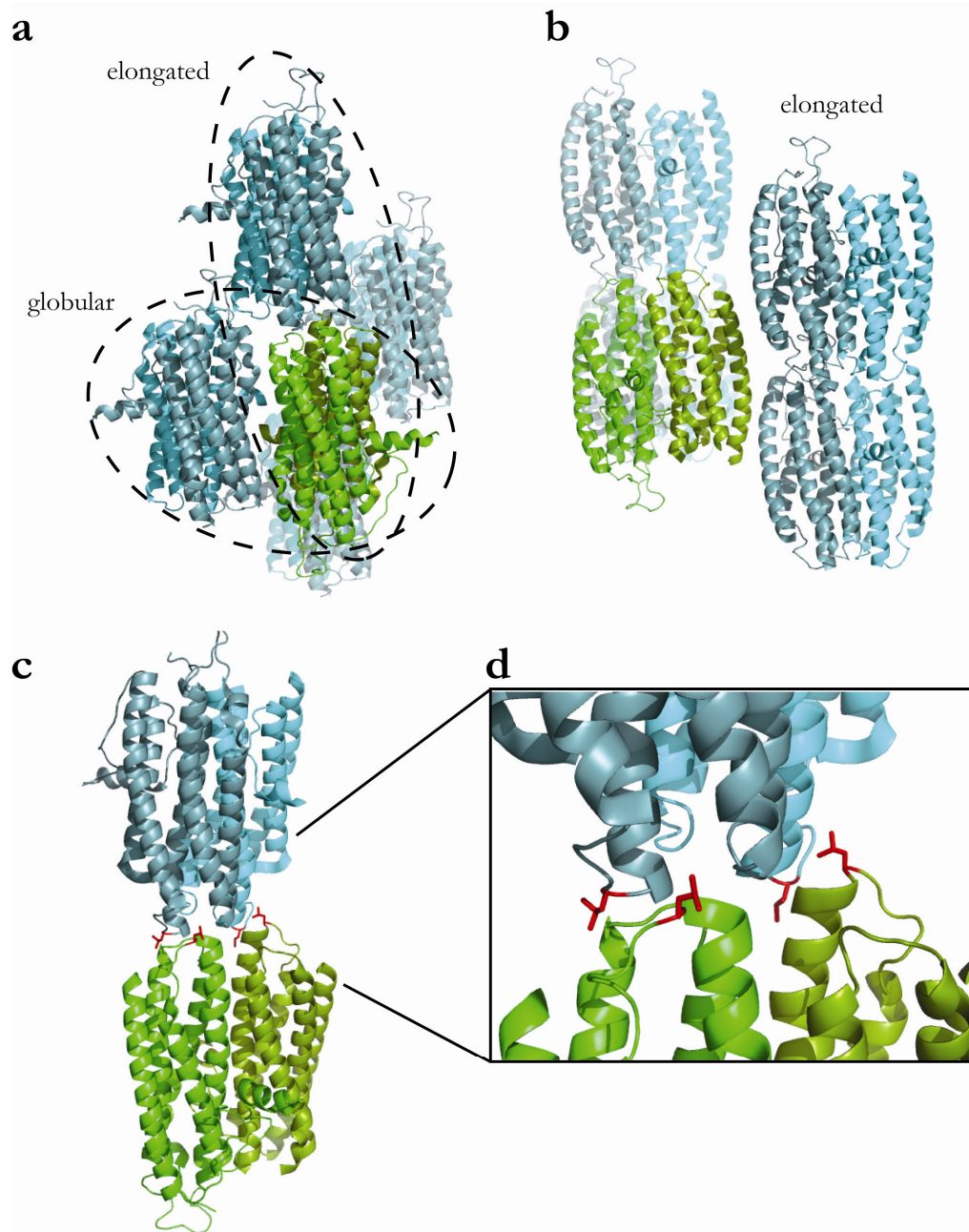


Figure 29: Crystallographic contacts of She2p. **a)** In the published crystal structure (PDB ID: 1XLY) a stable She2p dimer (green) was observed with a large buried surface interface of about 2015 Å²/monomer. Among the several crystallographic contacts (blue) only two dimer contacts have significant buried surface interfaces. One interaction involves the basic-helical hairpin-containing side on two She2p dimers, resulting in an upside-down orientation and was termed “globular tetramer”. The other interaction occurs through the uncharged upper surface in a head-to-head conformation and was termed “elongated tetramer”. **b)** Structural arrangement from a) rotated by roughly 90°. On the right side of the green dimer, weak crystallographic interactions with other dimers are visible. These contacts fail to bury significant surface areas and are unlikely to be relevant for oligomerization in solution **c)** The elongated tetramer with amino acid leucine 130 is shown with its side chains in red. Leucine 130 is located at the dimer-dimer interface. **d)** Close-up of the tetramer-interaction surface shown in c).

RESULTS

Subsequently, I analyzed if one of these crystallographic tetramers is consistent with the experimental SAXS-scattering curves of wild-type She2p. Using CRY SOL (Svergun et al., 1995) I calculated the theoretical scattering curves of the She2p dimer and of the two potential tetramers and compared them with the experimental curve of wild-type She2p at intermediate protein concentration (2.8 mg/ml) (Figure 30). The theoretical scattering curve of the elongated She2p tetramer fits reasonably well to the experimental data of wild-type She2p (Figure 30 a). In contrast, the theoretical scattering curve of the She2p dimer fit to a lesser extent and the curve of the globular tetramer differs substantially from the experimental data (Figure 30 b). Thus, formation of a globular tetramer can be excluded. This finding is consistent with previous mutational analysis of the lateral crystallographic contacts, for which no effect on oligomerization could be observed (Dierk Niessing, personal communication). In summary, it is unlikely that She2p exists preferentially as a dimer or as a globular tetramer, but it can be concluded that the elongated tetramer is the preferential assembly in solution.

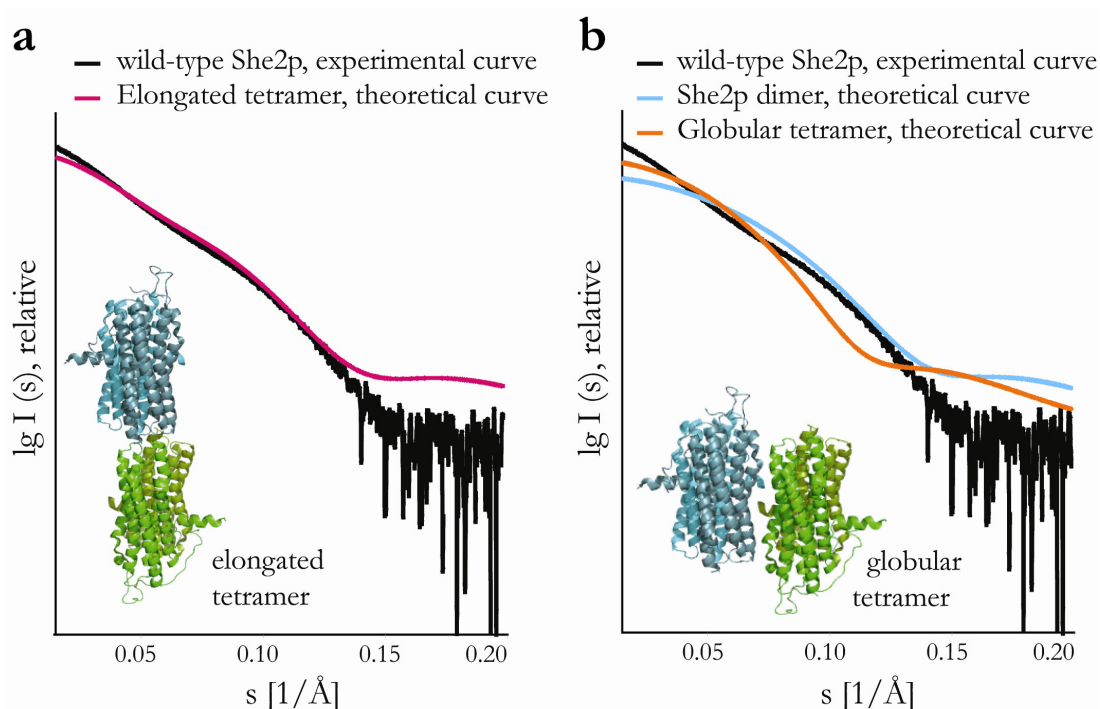


Figure 30: CRY SOL analysis suggests an elongated She2p tetramer in solution. a) The theoretical scattering curve of the elongated She2p tetramer (see inset) shows a good fit with the experimental scattering data of wild-type She2p at a concentration of 2.8 mg/ml. **b)** In contrast, neither the theoretical scattering curve of the globular She2p tetramer (see inset) nor the theoretical curve of a She2p dimer matches the experimental curve of She2p. Both tetramer models were generated from the She2p crystal structure as shown in Figure 29.

2.2.18 Low resolution models of wild-type She2p

Solution structures of wild-type She2p were determined with the *ab initio* dummy residue approach using the program GASBOR (Svergun et al., 2001). GASBOR compares thousands of configurations of a chain of dummy residues to the measured scattering curve. Dummy residues represent a string of beads of the size of an average amino acid. The number of dummy residues used for data fitting is equivalent to the number of amino acids in a given protein. Therefore it is very useful to know the oligomerization state of the measured sample. GASBOR is always run in at least ten cycles to generate one bead model per cycle. The resulting models are aligned and averaged to a final shape using the programs SUPCOMP and DAMAVER (Volkov and Svergun, 2003).

In order to produce reasonable models, a number of input files with different constraints were generated from the wild-type She2p SAXS data. Because a She2p tetramer with a potential symmetry axis was expected, different solution structures were calculated assuming either no symmetry (P1) or a two-fold symmetry (P2). The bead models were calculated by two different GASBOR programs, which fit the dummy residues to either the $P(r)$ function (GASBORp) or to the non-transformed scattering intensity I (GASBORi) and might produce different models. All GASBOR runs were repeated 10 times with identical parameters and the resulting bead models were aligned and averaged.

Only bead models that were calculated without symmetry constraints (P1) from data measured at low and intermediate concentration fit to the experimentally determined maximum dimensions of She2p (Table 6). In order to determine, which of the two possible tetrameric assemblies of She2p is present in solution, both crystallographic tetramers were overlaid with different *ab initio* solution structures using the program SUPCOMB (Figure 31). The elongated She2p tetramer fits quite reasonable to the dimensions of the SAXS solution structures with a length of approximately 145 Å and a width of approximately 50 Å (Figure 31). This corresponds roughly to the length of two She2p dimers (135 Å) with the constant width of one She2p dimer and thus, strongly suggests that the observed elongated tetramer is indeed the preferential She2p conformation in solution. However, apparently all solution structures show a specific bend in the middle of the particle, which is not consistent with the straight conformation of the crystallographic tetramer (Figure 31). In addition, significant parts of the crystal structure stick out of the bead models. This finding suggests that either the observed crystallographic contacts were enforced by the crystallization process or that the SAXS scattering data did not provide sufficient information for fully reasonable *ab initio* modeling.

RESULTS

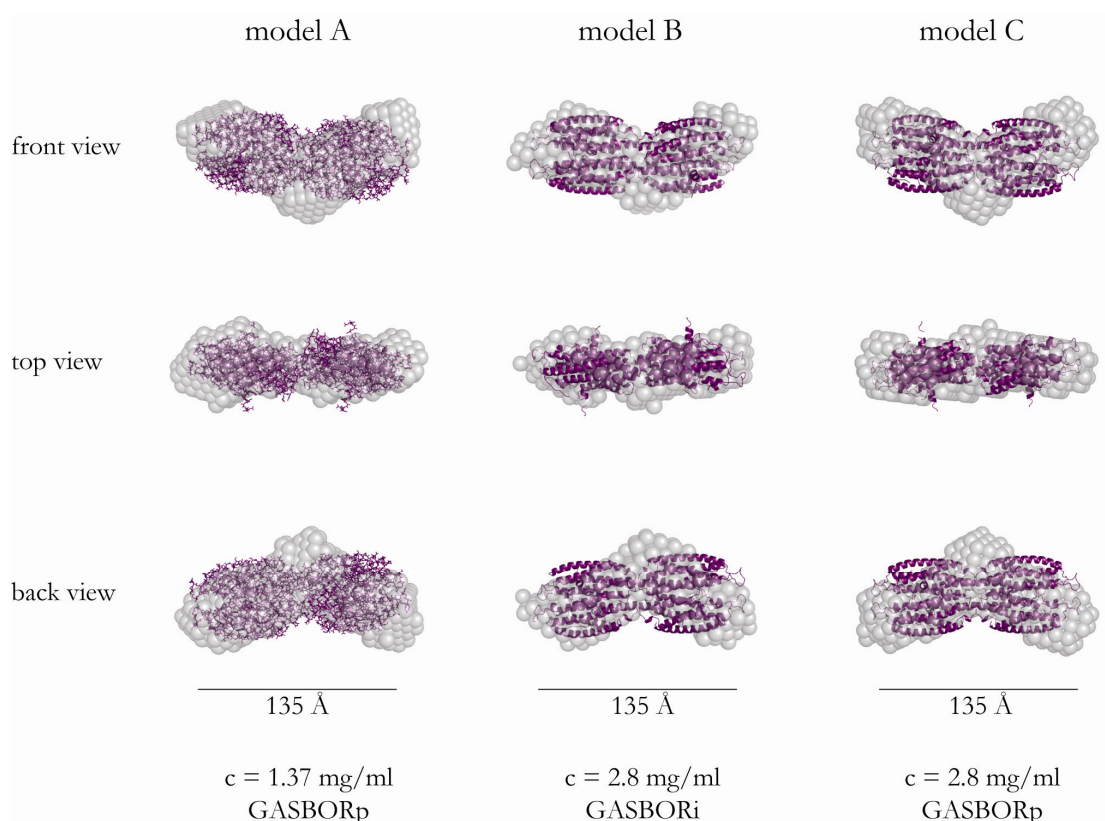


Figure 31: *Ab initio* modeling indicates an elongated tetramer of wild-type She2p. Depicted are three selected *ab initio* models of SAXS-solution structures of wild-type She2p. Bead models were calculated using either GASBORp or GASBORi from scattering data measured at the indicated protein concentration. For each model, 10 independent structures were aligned and averaged. The final bead models were aligned to the elongated She2p tetramer, which was derived from the She2p crystal structure. The scale bar represents 135 Å and matches roughly the length of the crystallographic She2p tetramer. In total, more than 20 different SAXS-solution structures were calculated and all of them exhibited a similar dimension and overall shape, yet never perfectly matched the crystallographic elongated tetramer.

Taken together, the overall dimension and the shape of various *ab initio* solution structures agree well with the elongated She2p tetramer that was modeled according to crystallographic contacts. Thus, a globular She2p tetramer can be excluded. Furthermore, the elongated tetramer forms most likely through the uncharged upper surface region of She2p. This conclusion is supported by the observed oligomerization defect of the mutant She2p-L130Y and the hence resulting impaired RNA binding *in vitro* and mRNP assembly *in vivo*.

Unfortunately, attempts to obtain reasonable solution structures of the She2p-L130Y mutant were not successful. However, I observed that She2p-L130Y forms smaller particles of a rather globular shape with an extended region on one side of the particle (not shown). Such an extension is consistent with the observation from the Kratky plot.

RESULTS

In conclusion, low-resolution *ab initio* models could be calculated from SAXS experiments. The bead models confirm all observations that She2p homo-oligomerizes to complexes larger than the reported dimer from the crystal structure. Most likely, tetramers are the physiological oligomerization state. However, in different buffer compositions She2p was observed to form larger particles than a tetramer. Thus, it cannot be excluded that higher-order oligomers might be important for mRNA binding and mRNP assembly *in vivo*, but She2p tetramers are the likely building blocks for such potentially larger complexes.

2.2.19 Attempt to obtain an atomic model of tetrameric She2p

As described above, the SAXS-solution structures do not perfectly match the elongated She2p tetramer, which was modeled according to existing crystallographic contacts (Figure 31). Therefore, I attempted to obtain a high-resolution structure by X-ray crystallography of She2p in its native oligomerization state. For crystallization experiments the variant She2p- Δ helixE was used. Deletion of helix E was shown not to affect oligomerization of She2p (chapter 2.2.8) and is believed to favor crystallization (chapter 2.1.2).

She2p- Δ helixE was crystallized in 96-well plates with a total precipitant volume of 50 μ l by combining 0.5 μ l protein solution (4 mg/ml) and 0.5 μ l precipitant solution. Initial protein crystals were obtained in conditions #A9 (0.1 M Bis-Tris pH 5.5, 3 M NaCl), #C11 (1 M ammonium sulfate, 0.1 M HEPES pH 7.0, 0.5 % (w/v) PEG 8000) and #C12 (15 % (v/v) Tacsimate pH 7.0, 0.1 M HEPES pH 7.0, 2 % (w/v) PEG 3350) from the Hampton Index screen. Crystals obtained in #C11 could not be further optimized. Refining the #A9 growth conditions yielded as best results thick adherent plates of a maximum size of 40 μ m x 40 μ m x 5 μ m (Figure 32 a). Variation of #C12 resulted in growth of relatively large elliptic crystals with a maximum size of about 100 μ m x 25 μ m x 20 μ m (Figure 32 b). Several crystals derived from #A9 and #C12 were tested for diffraction at the ESRF, but no diffraction pattern was obtained for any of them. Further optimization of both crystal forms was not successful.

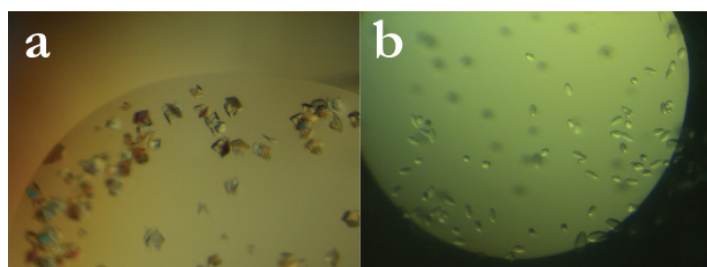


Figure 32: Crystals of She2p- Δ helixE.

Crystals of She2p- Δ helixE were obtained in optimized conditions **a)** 0.1 M HEPES pH 7, 3 M NaCl and **b)** 15 % (v/v) Tacsimate pH 7, 0.1 M HEPES pH 7, 1 % (w/v) PEG 3350. Both crystal forms showed no X-ray diffraction.

3 Discussion

mRNAs destined for subcellular localization are transported as part of large messenger ribonucleoprotein particles (mRNPs). These mRNPs usually contain a number of proteins that mainly function in mRNA binding, translational silencing of the mRNA, and active transport of the particle (Martin and Ephrussi, 2009). The most comprehensively understood mRNA-transport complex is the *ASH1* mRNP from *Saccharomyces cerevisiae*. It comprises *ASH1* mRNA and the three protein-core factors She2p, She3p, and Myo4p (Böhl et al., 2000; Jansen et al., 1996; Long et al., 2000; Takizawa and Vale, 2000). Recent studies suggested that besides *ASH1* mRNA more than 23 additional mRNAs are transported in a She2p-dependent manner (Oeffinger et al., 2007; Shepard et al., 2003; Takizawa and Vale, 2000). The RNA-binding protein She2p is assumed to be responsible for specific incorporation of target mRNAs into the translocation complex (Böhl et al., 2000; Kruse et al., 2002; Long et al., 2000), but thus far, the molecular mechanisms underlying She2p-dependent mRNP assembly are poorly understood. In this study, the requirements for specific incorporation of target mRNA into a functional mRNP were investigated.

3.1 She2p binds RNA through a composite binding site

RNA-binding studies revealed that She2p binds to the zipcodes of *ASH1*, *EAR1*, and *WSC2* with affinities ranging between 100 nM (*ASH1* E3) and 770 nM (*EAR1*) (Figure 10). Earlier it was suggested from Electrophoretic Mobility Shift Assays that She2p binds apparently with varying affinities to the respective *ASH1* zipcodes (Böhl et al., 2000; Long et al., 2000), thus corroborating the differences in affinity I have measured. The observed affinities are comparable with a binding constant of $K_D=200$ nM that was measured for the interaction of the neuronal Fragile X-mental Retardation Protein (FRMP) with one of its targets, the non-coding RNA BC1 (Zalfa et al., 2005), suggesting a similar range of affinity. In contrast, the chick Zipcode Binding Protein 1 (ZBP1) binds remarkably stronger to its target mRNA *β -actin* with a $K_D=10$ nM (Farina et al., 2003). However, the above mentioned RNA:protein complexes are two of the rare examples for which binding constants are known. Most RNA:protein interactions within mRNA-localization complexes are not quantified to date. Therefore, it is not clear what affinities are required for efficient mRNA transport. Furthermore, affinity does not necessarily correlate with complex stability, which might

DISCUSSION

suggest that for She2p-dependent mRNA transport additional factors are needed to stabilize the She2p:mRNA complex and to compensate for a lower affinity (see chapter 3.4).

Surprisingly, She2p can also bind to a subset of unrelated stem-loop containing RNAs with an affinity in the lower micromolar range (Figure 10). In contrast, no binding was measured to an unstructured poly(A) RNA, suggesting that She2p preferentially recognizes double-stranded RNA (dsRNA) or single-stranded regions in context of a stem-loop structure, which are well proven features of RNA-localization elements (Jambhekar and DeRisi, 2007). Apparently, She2p is able to discriminate between target and non-target RNAs and raised the question, which factors mediate specific interactions between She2p and the number of mRNA cargoes in yeast.

She2p was proposed to specifically recognize a degenerate RNA motif in various zipcodes of bud-localizing mRNAs (Jambhekar et al., 2005; Olivier et al., 2005). Sequence-specific RNA recognition is a common feature of RNA-binding proteins that have well-defined RNA-binding motifs like double-stranded RNA-binding domains (dsRBDs), RNA-recognition motifs (RRMs), hnRNP K-homology (KH) domains, or the Pumilio/FBF (PUF) domain (Auweter et al., 2006; Lunde et al., 2007). RNA-binding proteins implicated in mRNA localization, like *Drosophila* Staufen, chick ZBP1, or human Pumilio, usually contain at least one type of the above mentioned RNA-binding motifs (Farina et al., 2003; St Johnston et al., 1992; Zamore et al., 1997). However, She2p lacks any homology to known RNA-binding proteins and its RNA-binding surface, though rich in conserved arginine residues could not be attributed to any class of RNA-binding motifs (Niessing et al., 2004). Investigation of She2p's binding to short RNA-consensus sequences, which were taken out of their structural context, revealed an only weak affinity compared to full-length zipcode binding (Figure 13). This finding demonstrates that RNA-sequence elements alone do not provide specificity for She2p binding. It is rather likely that She2p recognizes the degenerate RNA motifs in context of secondary structures as corroborated by She2p's affinity to stem-loop structures in general.

By mutational analysis, several conserved regions in She2p were identified, to which distinct RNA-binding features of She2p could be attributed. Firstly, the positively charged basic helical hairpin motif was shown to provide a platform for basal RNA binding, since mutations in the basic helical hairpin (N36S and R63K) resulted in a complete loss of RNA binding *in vitro* (Figure 15). This result is consistent with previous studies, in which *ASH1*-E3 binding *in vitro* and *ASH1*-mRNA localization *in vivo* was abolished upon mutation of single amino acids in the basic helical hairpin (Gonsalvez et al., 2003; Niessing et al., 2004). In contrast to

DISCUSSION

observations made by Gonsalvez et al. (2003), She2p becomes trapped in the nucleolus when it completely fails to bind RNA (Du et al., 2008). Such a nucleolar accumulation was also reported for the mammalian Staufen 2 protein, which is implicated in neuronal mRNA localization and becomes enriched in nucleoli upon mutation of its RNA-binding domain (Macchi et al., 2004). Since the double mutant She2p-N36S-R63K is also unable to bind to non-specific RNAs (Figure 15), it can be assumed that She2p's nucleolar entrapment is caused by the loss of RNA binding and not by any unspecific interaction with nucleolar RNA species such as rRNAs or snoRNAs. Along with a previous study (Kruse et al., 2002), this finding confirms that She2p is exported to the cytoplasm in an mRNA-dependent fashion. Thus, mRNA binding by She2p might be a prerequisite for export of She2p and further indicates that the assembly of a functional mRNP begins in the nucleus with the selection of the mRNA cargo (see chapter 3.4).

Secondly, the small protruding helix E and the very C-terminus of She2p were identified to be required for specific recognition of individual *ASH1* zipcodes, but not for binding to unrelated stem-loop structures *in vitro* (Figures 17 and 18). Respective She2p mutants fail to assemble into She3p-containing transport complexes and thus, do not localize to the bud tip *in vivo*. This strongly suggests that She2p selects its target mRNA *ASH1* through two conserved regions, the helix E and the C-terminus, and by this means, discriminates between specific bud-localizing mRNA and non-specific RNA. Importantly, the observed failure in mRNP assembly indicates that functional transport complexes only assemble when the correct target mRNA is bound by She2p, supporting an mRNA-dependent complex assembly. This is confirmed by previous studies in which an *ASH1* mRNA-dependent association of She2p with the cytoplasmic She3p-Myo4p complex was postulated (Böhl et al., 2000; Long et al., 2000).

As mentioned above the RNA-binding mutant She2p-N36-R63K accumulates in the nucleolus and is not exported to the cytoplasm. However, immunostaining experiments showed cytoplasmic localization of She2p variants with impaired *ASH1*-mRNA binding (She2p-ΔhelixE and She2p-ΔC-term) (Figure 20). This raised the question why these mutants were not enriched in the nucleus although RNA binding is impaired. Two explanations are possible. Firstly, *ASH1*-mRNA binding by She2p-ΔhelixE and She2p-ΔC-term was only reduced and unrelated RNAs could be bound with unaltered affinity. In contrast, RNA binding by the basic helical hairpin mutant was completely abolished. Since both helix E and the C-terminus specifically recognize *ASH1* mRNA, one RNA-binding region might be able

DISCUSSION

to compensate for the loss of the second RNA-binding surface. As a consequence, reduced, but specific RNA binding may be sufficient for mRNA-mediated export. The second possibility is that the mutants She2p- Δ helixE and She2p- Δ C-term are never imported into the nucleus. Nuclear import of She2p was shown to depend on the importin α Srp1p (Shen et al., 2009). For this, She2p interacts with Srp1p through a non-classical nuclear localization signal that is located on the side of the She2p dimer (Figure 4). Since the deletion of helix E or the C-terminus interferes in no way with the NLS, both She2p mutants are likely to be imported into the nucleus. However, in order to address this question thoroughly, nuclear She2p localization could be assessed by immunostaining experiments against She2p mutants upon block of mRNA export in a *mex67-5^{ts}* strain (Hurt et al., 2000). Enrichment of She2p in the nucleus would be a clear indication for proper She2p import.

Because deletion of helix E influenced only binding to *EAR1* mRNA and deletion of the C-terminus had no effect on either *EAR1*-mRNA or *WSC2*-mRNA binding, these two bud-localizing mRNAs are likely to be recognized by She2p in a different way than *ASH1* mRNA, (Figures 17 and 18). This finding suggests that besides *ASH1*-like mRNAs there is at least one other class of RNA, which is bound by She2p in a different manner. This hypothesis is supported by a recent publication, in which the PUF protein Puf-11 from *C. elegans* was shown to recognize three different classes of RNAs by using multiple binding modes (see also chapter 3.5) (Koh et al., 2009). However, it remains unclear how specific She2p binding to these mRNAs is achieved. Nevertheless, selective disruption of *ASH1*-specific recognition motifs in She2p (helix E and C-terminus) completely abolishes She2p localization and mRNP formation *in vivo*. Thus, it affects all She2p-dependent mRNAs.

The identified regions required for RNA binding cover large surface areas on both sides of the She2p dimer (Figure 33 a, b). Together, the three individual binding motifs form a composite binding site, each consisting of the positively charged basic helical hairpin motif, the small protruding α -helix E, and the very C-terminus that was not visible in the She2p-crystal structure and is most likely unstructured (Figure 33 a, b) (Niessing et al., 2004). In order to find out if this composite binding region is able to accommodate double-stranded RNA, the missing C-terminus was modeled by fusing an unstructured loop of nine alanine residues to amino acid 237 of the crystallographic model (Figure 33 a-c). Assuming that the C-terminus is flexible, the regions required for RNA binding form a gripper-like surface feature with a cleft of about 23 Å to 25 Å in diameter. Double-stranded RNA has an average diameter of approximately 20 Å. Thus, I conclude that zipcode RNAs fit into the modeled RNA-binding

DISCUSSION

cleft, even taking into consideration the structural variability of zipcodes having unpaired nucleotides in context of double-stranded regions. To visualize this conclusion, a random double stranded RNA was taken from the Protein Data Bank (PDB) and manually docked into the binding cleft. As shown in Figure 33 c, the stem loop of the RNase P ribozyme P4 domain (27 bases; PDB-ID: 2CD1) indeed fits into the cleft and is able to contact all three RNA-binding motifs. In analogy to a gripper, the binding cleft could accommodate RNAs of different conformations, which could directly contact the protruding helix E and the basic helical hairpin on one side. In addition, cooperative binding by the flexible C-terminus could account for different tertiary structures of the respective zipcode RNA. In summary, the individual binding motifs in She2p form a distinct RNA-binding cleft on each side of the She2p homodimer. This composite binding site is required for specific recognition of *ASH1*-like mRNAs.

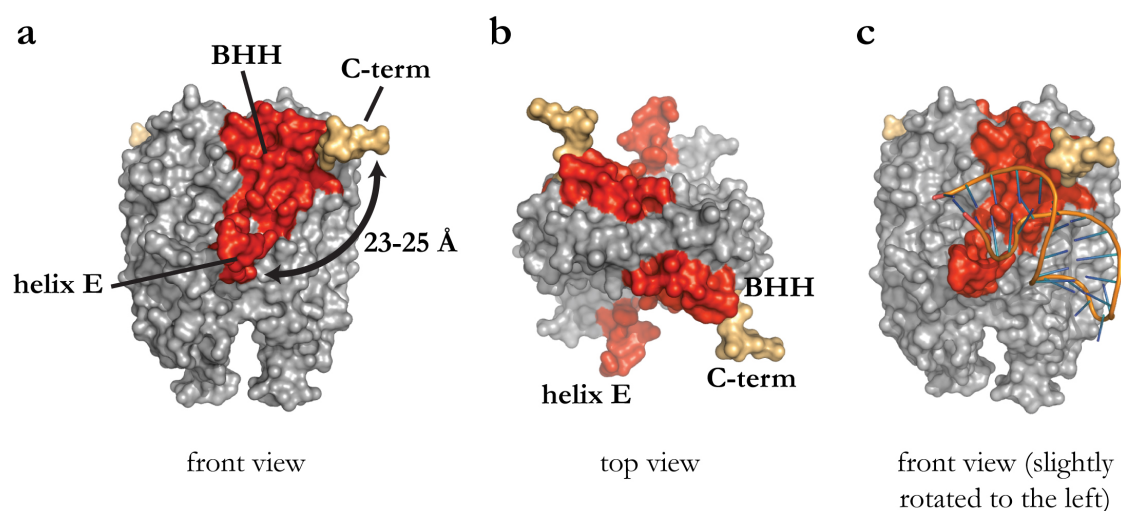


Figure 33: Composite RNA-binding site in She2p. a)-c) Depicted is a model of the composite gripper-like RNA-binding region in a She2p dimer. Regions required for RNA binding are marked in red (helix E and the basic helical hairpin (BHH)). The C-terminus (C-term), for which no structural information is available, was modeled with an unstructured nine-alanine loop and is labeled in beige. a) shows She2p in front view, b) in top view, and c) in front view with a slight rotation to the left. The composite binding motif forms a cleft with a diameter of 23-25 Å and is large enough to accommodate double-stranded RNA. c) The RNase P ribozyme P4 domain was manually aligned onto the basic helical hairpin motif between helix E and the modeled, unstructured She2p C-terminus.

In order to explain how She2p exactly binds to several target mRNAs, three-dimensional structures at atomic resolution are required. In this study, She2p was crystallized with three different stem-loop RNAs that were derived from the *ASH1*-E2B zipcode. Unfortunately, the crystals of the different She2p:RNA co-complexes diffracted X-rays to maximum resolution

of only 9 Å, which was not sufficient to obtain any structural information. Crystallization of respective She2p:RNA co-complexes turned out to be very challenging, since the complex RNA-binding mechanism by She2p significantly restricted the number of potential She2p constructs for crystallization (chapter 2.1.2). In addition, also the number of RNA targets was limited due to the length of the different zipcodes (up to 120 nucleotides), which made a chemical synthesis impossible. This technical problem could in future be circumvented for potential crystallization studies by using the tRNA-scaffold approach (Ponchon and Dardel, 2007). This technique enables the production of longer RNAs in high amount by *in vivo* transcription in *E. coli* cells. During this study, I established this method and indeed, obtained in a tentative experiment several milligrams of different zipcodes, which can be purified using standard size-exclusion chromatography (data not shown). Thus, the tRNA-scaffold approach represents a powerful method for the generation of RNA in large scale and may help to overcome technical limitations in chemical synthesis of RNAs. So far, this technique was not applied for crystallization studies with She2p and *ASH1* RNAs, but may provide the basis for further experiments. However, a crucial point for successful crystallization of an She2p:RNA complex is the complex stability. In RNA-binding experiments, I observed that complexes formed between She2p and different *ASH1* zipcodes are rather transient, even though the RNA is bound with high affinity. The low complex stability is very likely to account for the difficulties in crystallizing She2p with RNA. One possibility could be to test longer RNAs produced by the tRNA-scaffold method that might bind to a larger surface area in She2p than shorter hairpin RNAs do. The best possibility, however, for obtaining a stable complex is to crystallize a trimeric complex consisting of RNA, She2p, and the adapter protein She3p. For this purpose, She3p has to be expressed in insect cells or yeast cells, because, at least in my hands, the She2p-interacting region of She3p cannot be produced by recombinant expression in *E. coli* cells (data not shown). Taken together, considerable effort will have to be made to obtain insights into the interaction of She2p with one of its mRNA cargoes at the atomic level.

The observation that She2p binds RNA through a large surface area is supported by the fact that many RNA-binding proteins have a modular structure (Lunde et al., 2007). They are composed of multiple RNA-binding domains (see above), which are presented in various structural arrangements to expand the functional landscape of the protein (Burd and Dreyfuss, 1994). These domains act in a cooperative fashion in order to accommodate the structural diversity of the RNA substrates. As a consequence, such a concerted effort results in RNA recognition with a specificity and affinity that would not be achieved by a single RNA-binding

domain. Specific RNA recognition is often mediated by two binding motifs in tandem, which are connected by a linker and undergo extensive interdomain arrangements upon RNA binding, thus providing an extended RNA-binding surface. For instance, this mechanism has been observed for the two RRM of the yeast Hrp1 protein as well as for the two KH domains in the transcription factor NusA (Beuth et al., 2005; Perez-Canadillas, 2006; Worbs et al., 2001). Such an RNA-dependent conformational change could in principle also apply to the flexible C-terminus of She2p in order to deal with different structural conformations of various zipcodes. Another prominent example of a modular domain protein is Staufen. Staufen comprises five dsRBDs, three of which (dsRBDs 1, 3, and 4) bind to double-stranded RNA (Micklem et al., 2000; St Johnston et al., 1992). Interestingly, the interaction of isolated dsRBDs with dsRNA is very weak (μM or weaker) (Ramos et al., 1999; Ramos et al., 2000), suggesting that the target mRNAs are bound by the three dsRBDs in a concerted manner or that additional stabilizing factors join the mRNA:protein complex. However, it remains to be proven if the three domains act cooperatively *in vivo*, since the individual dsRBDs bind to RNA *in vitro* and, thus far, only the RNA-binding activity of dsRBD 3 was shown to be essential for the interaction of Staufen with *oskar* and *bicoid* mRNA *in vivo* (Ramos et al., 2000). Besides Staufen, also *Drosophila* Bruno binds to *oskar* mRNA, yet to distinct Bruno Response Elements (BREs) and thereby represses translation of non-localized *oskar* (Chekulaeva et al., 2006; Kim-Ha et al., 1995; Nakamura et al., 2004; Webster et al., 1997; Wilhelm et al., 2003). Bruno contains three RRMs, which form two distinct RNA-binding modules (Snee et al., 2008; Webster et al., 1997). Similar to helix E and the C-terminus in She2p, each binding domain of Bruno can bind to BRE-containing RNA, but both are required for maximal binding *in vitro* and full translational repression activity *in vivo* (Snee et al., 2008). One of the most sophisticated mechanisms of cooperative binding, however, is represented by the Pumilio domains, which are arranged in tandem repeats (Zamore et al., 1997). Each domain can recognize a single nucleotide on its own, but by combining multiple repeats, up to eight nucleotides can be specifically bound with high affinity (Wang et al., 2002). Thus, the Pumilio-repeat proteins likely provide the most extreme example of an extended composite RNA-binding site.

3.2 She2p tetramers are the functional units

The crystal structure of She2p as well as previous analytical ultracentrifugation experiments revealed that She2p forms homo-dimers (Niessing et al., 2004). However, these studies were performed with a variant of She2p (She2p-(6-239)- Δ Cys; see chapter 2.1.2) that was created to facilitate crystallization. In my hands, this variant has lost its affinity to short degenerate RNA motifs (chapter 2.1.6) most likely due to the C-terminal truncation. Since *ASH1* binding strongly relies on a functional C-terminus (chapter 2.2.9), I conclude that also *ASH1*-zipcode binding is disrupted in this mutant. In addition I observed that truncation of the C-terminus apparently impairs the oligomeric state of She2p (Figure 18 b). In summary, it has to be assumed that the crystallized She2p might not represent the functionally active form *in vivo*.

In order to investigate the oligomeric state of She2p, several experiments were performed in this study. Analytical ultracentrifugation (AUC) studies revealed that She2p forms large particles of approximately 290 kDa, corresponding to a decameric complex (Figure 24). For symmetry constraints, however, also octameric or dodecameric complexes would be possible. Identification of the variant She2p-L130Y led to the conclusion that She2p assembles into higher-order oligomers through the upper uncharged surface region, because this mutant protein failed to form wild-type like complexes (Figure 24). However, the AUC experiments were performed at protein concentrations that are far above the physiological concentration of She2p *in vivo* (230 nM; Niessing et al., 2004). Furthermore, She2p turned out to assemble into large particles dependent on the protein concentration and the buffer composition as judged by Dynamic Light Scattering (Tables 3 and 4). Therefore, I determined the oligomeric state of She2p at protein concentrations below and above the calculated physiological concentration of 230 nM using analytical size-exclusion chromatography. I found that She2p is a stable tetramer above and, importantly, even below the physiological protein level (Figure 25). Analysis of the She2p-L130Y mutant revealed a clear dimer with the tendency to form larger particles at higher protein concentrations, thus confirming that She2p oligomerizes through its upper surface. From these results I conclude the following: Firstly, the crystallographic dimer most likely does not represent the physiological oligomerization state of She2p. The dimers are rather subunits of the identified She2p tetramer. Secondly, the large complexes observed in AUC experiments probably result from a high protein concentration. Thirdly, She2p tetramers are presumably the functional units *in vivo*, since mutation of the upper surface abolishes She2p localization and mRNP assembly (Figures 20 and 22, respectively). The observed

DISCUSSION

reduction in RNA binding *in vitro* (Figure 16) can be explained to result directly from disruption of the oligomeric state.

SAXS analysis confirmed that She2p forms tetrameric complexes of an elongated shape, most likely through the upper surface area. Analysis of crystallographic dimer contacts revealed two possible tetrameric assemblies, resulting in either a globular tetramer or an elongated tetramer (Figure 29). The calculated scattering curve of the elongated tetramer matches well with the experimental scattering curve of She2p, indicating an elongated shape (Figure 30). This is further supported by the $P(r)$ function of She2p, which suggests a stretched shape (Figure 28). In contrast, the She2p-L130Y mutant formed smaller molecules of a rather globular shape, indicating that the elongated tetramer was disrupted (Figure 28). Such a disruption can be explained by the mutation of the small hydrophobic leucine 130 into the large tyrosine that results in a steric clash in the dimer interface and prevents the association of the complementary dimer. The constantly observed concentration-dependent oligomerization of She2p-L130Y is easily explained by the fact that only one residue in the dimer interface was mutated. Such a single mutation is apparently not sufficient to completely disrupt oligomer formation, which is apparently forced at higher protein concentrations. In order to achieve a complete dissociation of the tetramer into dimeric subunits, mutations of additional residues within the dimer interface would be required. A promising candidate to analyze is threonine 47. Next to leucine 130, threonine 47 is located at the upper surface region (Figure 34) and its mutation into tyrosine is supposed to disrupt She2p oligomerization as it was observed in size-exclusion chromatography (Dierk Niessing, personal communication).

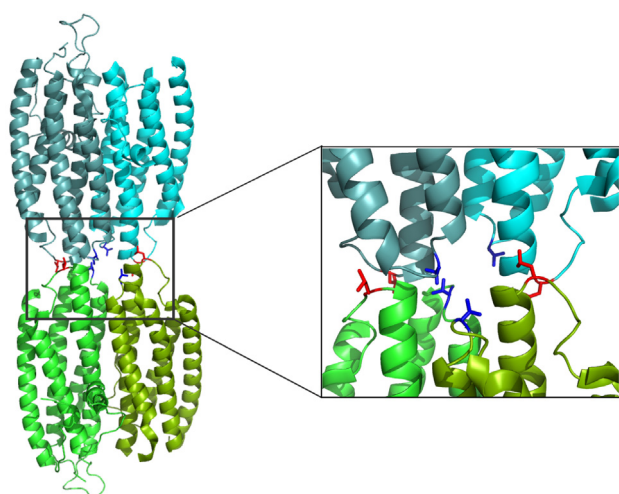


Figure 34: Contacts in the upper surface region of She2p. Depicted is the elongated tetramer, which was derived from the She2p-crystal structure (compare with Figure 29). Amino acids in the dimer interface that interfere with She2p-tetramer formation upon mutation are shown. Leucine 130 is shown with its side chains in red, threonine 47 is shown in blue.

DISCUSSION

Thus, She2p is likely to form an elongated tetramer through an asymmetric head-to-head interaction, for which an atomic model can be derived (Figures 35, see also Figure 29). In this model, the different RNA-binding motifs form a continuous binding surface on both sides of the molecule. Thus, an extended surface area is provided in order to deal with the number of structural variations of different bud-localizing mRNAs. Since the exact conformation and orientation of the She2p C-terminus is unknown, it was omitted from the tetrameric model. However, the flexible C-terminal region of She2p might change its conformation in response to the respective mRNA cargo that binds to She2p. This assumption would also be consistent with the observation that the bud-localizing mRNAs *EAR1* and *WSC2* are not recognized by the C-terminus of She2p and are apparently bound in a different manner than *ASH1* mRNA.

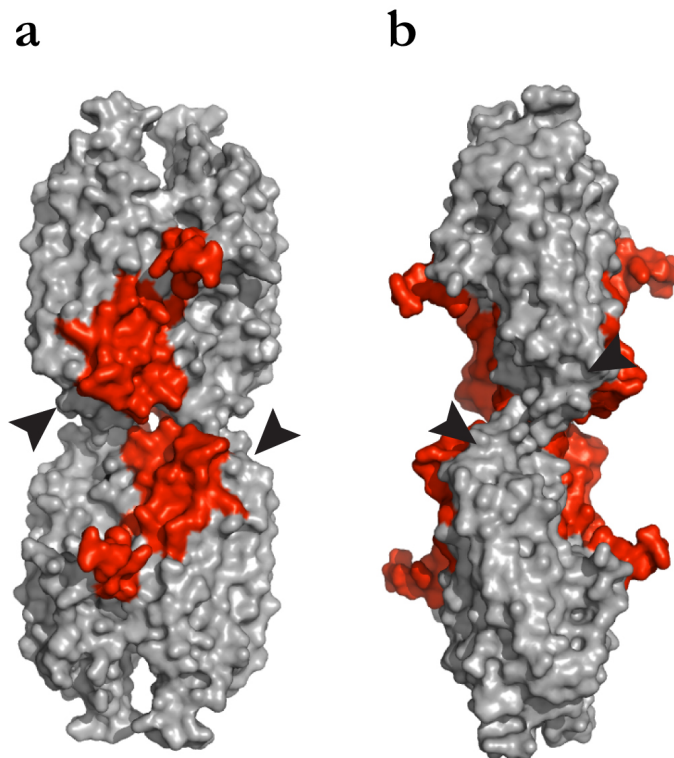


Figure 35: She2p tetramers provide an extended RNA-binding region.

Depicted is a surface representation of the elongated tetramer from the front (a) and from the side (b, rotated by about 90°). The RNA-binding motifs basic-helical hairpin and helix E are highlighted in red. For a better overview, the modeled C-terminus of She2p from Figure 33 is not shown. Instead, its position is indicated on two monomers each per tetramer by arrow heads.

In principle, also a tail-to-tail or head-to-tail orientation of two She2p dimers would result in an elongated shape of the tetramer. Since mutation of the upper surface region disrupts the oligomeric state of the tetramer, the tail-to-tail orientation can be excluded. A head-to-tail association would still involve the upper surface of one dimer, but also the less conserved loop-containing bottom region of She2p. In this model, mutation of leucine 130 to tyrosine could lead to disruption of the tetramer, which in turn would be also assumed upon deletion of the bottom loops. However, when I purified a respective She2p mutant, which was

DISCUSSION

intended for crystallization (She2p- Δ loop, chapter 2.1.2), I did not observe any oligomerization defect, rendering this type of association unlikely. Likewise important, in a head-to-tail orientation no extended RNA-binding surface would be provided. For these reasons, I can exclude both a tail-to-tail and a head-to-tail association and refer to the head-to-head-orientation as the most likely model.

As shown in Figure 35, the extended RNA-binding surface of the She2p tetramer forms only when the two dimeric subunits are skewed to each other as a direct result from the crystallographic contacts. One can only speculate if the orientation of the crystallographic tetramer is entirely correct or if the observed contacts were enforced during the crystallization process. This example shows that it is meaningful to prove, if the obtained crystal structure indeed represents the native conformation. For this purpose, SAXS can be used to calculate *ab initio* models of the protein and compare them with the crystal structure. For She2p, a number of different *ab initio* models were calculated that all roughly matched the shape of the crystallographic elongated tetramer (Figure 31).

The overall dimension of the SAXS-solution structures was consistent with the one of the elongated tetramer. However, all *ab initio* models exhibited a significant bend in their central region and thus, differed from the straight conformation of the elongated tetramer. One possible explanation is that crystal packing and the lack of the C-terminus in the crystal structure modified the exact conformation of the crystallographic tetramer, implying that the model determined in solution is correct. The second possibility is that the SAXS data did not provide sufficient information for exact calculation of the shape of the molecule. This might be likely since the scattering curves that were measured at different protein concentrations were not completely identical. Regardless which of both possibilities is true, the overall conclusion that She2p forms an elongated tetramer is beyond all question. Interestingly, calculations with symmetry constraints (P2) did not produce any trustworthy models. This might be explained by the screwed conformation of the dimers to each other, which does not reflect a two-fold symmetry. The third possibility is that both the in solution model and the crystallographic tetramer is correct. In this case one could speculate that She2p forms a bent shape without an mRNA ligand, which is reflected by the SAXS model. Upon mRNA binding, She2p might change its conformation into an elongated shape with an extended RNA-binding surface, which was accidentally trapped in the crystal structure.

It is very important to carefully interpret the structural models obtained from SAXS measurements. SAXS scattering data represent the averaged data from all possible orientations

of the molecule in solution. In order to calculate a three-dimensional model *ab initio* from the two-dimensional scattering curve it is very useful to compare the resulting model with structural data that were obtained with a different technique, for instance X-ray crystallography. In case of the She2p tetramer, SAXS was helpful to dissect the potential tetrameric assemblies and by this means, to identify an elongated shape in solution. However, the *ab initio* models of the She2p tetramer did not perfectly match the elongated crystallographic tetramer. A significant disadvantage of SAXS is that no real parameter for quality control exists, which allows one to estimate the reliability of a calculated structure. In crystallography, the R-factor gives information about the significance of a structural solution. From SAXS data, however, a three-dimensional model can always be calculated. To summarize, SAXS is extremely powerful to obtain information of the solution structure of the protein when additional structural information of the protein are available. SAXS-solution structures calculated using the *ab initio* approach then have to be analyzed carefully for their reliability to make sure that no artifacts were generated.

3.3 She2p-tetramer formation allows simultaneous binding of different RNAs

She2p variants with a defect in oligomerization (She2p-L130Y and She2p-ΔC-term, respectively) also showed impaired RNA binding. This result suggests that specific RNAs interact with both dimer subunits simultaneously along the extended binding surface of the She2p tetramer. In total, a She2p tetramer has four gripper-like RNA-binding motifs (two on each dimer), which in turn form two large surface areas for RNA binding (Figure 35). Previous studies suggested for the interaction of a She2p dimer with *ASH1*-mRNA zipcodes a molar ratio of either 1:1 or 1:2, depending on the analyzed zipcode (Niessing et al., 2004; Olivier et al., 2005). When these molar ratios are translated into a She2p tetramer, one tetramer would bind to either two or four zipcode RNAs. Two zipcodes can be easily positioned on one tetramer when each zipcode is bound along the extended binding surface at each side of the tetramer. Binding of four zipcodes to one tetramer can be explained by the four gripper-like motifs binding to one zipcode each. The distance between the two protruding helices that confine the binding surfaces on each side of the tetramer is about 55 Å. Assuming an average diameter of a zipcode RNA of less than 25 Å, two zipcodes can be

DISCUSSION

accommodated next to each other without steric clashes on each side of the tetramer, resulting in four bound molecules.

These considerations suggest that zipcodes from different bud-localizing mRNAs can be bound by She2p simultaneously. This assumption is consistent with the observation that several different bud-localizing mRNAs are transported together in the same particle (Lange et al., 2008). For instance, *ASH1* has been shown to move to the bud tip together with *IST2* and *WSC2*, respectively. Due to technical limitation only pairs of mRNAs could be visualized in a single particle, but it is very likely that more than two mRNAs are transported simultaneously. This idea is supported by the fact that overexpression of *ASH1* mRNA results in the formation of fewer but larger *ASH1* mRNPs *in vivo* (Bertrand et al., 1998; Lange et al., 2008). Therefore, the formation of higher-order mRNPs by cross-linking individual She2p tetramers can be suggested. Such an interconnection of several She2p molecules could be achieved by association of She2p with mRNAs that contain multiple zipcode elements, like *ASH1* or *WSC2*. Thus, it is unlikely that She2p-dependent mRNPs have a defined size, but are rather heterogeneous in their RNA composition and dimension.

It is well proven that mRNPs implicated in mRNA localization contain an increasing number of mRNAs. For instance, the neuronal mRNP component FMRP associates with hundreds of mRNAs, whereas Staufen I, Staufen II and YB-1 are reported to co-purify with even thousands of RNAs (Darnell et al., 2005; Evdokimova et al., 2006; Furic et al., 2008). These findings do not necessarily imply that this vast amount of mRNAs is found in the same mRNP at the same time. It is rather likely that these mRNAs have been co-purified from a heterogeneous population of mRNPs, suggesting that the effective number of transcripts per particle is much smaller. However, at least a significant number of mRNAs will be transported in the same mRNP, which in turn leads to the assumption that the respective RNA-binding proteins are interconnected through their interaction with various mRNA cargoes.

The assembly of oligomeric mRNPs was also demonstrated for *oskar*-containing mRNPs in *Drosophila*. The splicing process of *oskar* mRNA is essential for the localization of *oskar* at the posterior pole of the oocyte (Hachet and Ephrussi, 2004). However, also intron-less reporter mRNAs bearing the *oskar* 3' UTR localized to the posterior by "hitch-hiking" with endogenous *oskar* mRNA, providing experimental evidence that two different RNAs are transported in the same particle (Hachet and Ephrussi, 2004). This 3' UTR-dependent mRNA oligomerization is mediated by the polypyrimidine tract binding protein (PTB) and is required for effective translational repression of *oskar* mRNA (Besse et al., 2009). Consistently, it was

shown that oligomerization of *oskar* mRNA depends on *Drosophila* Bruno and likewise results in the formation of large translationally silenced particles (Chekulaeva et al., 2006). Similarly, the formation of large She2p-dependent mRNPs might help to maintain a translational repressed state of all transported mRNAs, thereby supporting the previously described mechanisms for translational silencing of *ASH1* mRNA (Deng et al., 2008; Du et al., 2008; Paquin et al., 2007; Shen et al., 2009).

3.4 How are She2p-dependent priming complexes stabilized? - Implications for She2p-dependent mRNP assembly in *S. cerevisiae*

As described in chapter 3.1, She2p can bind to non-specific stem-loop containing RNAs. Surprisingly, the difference in affinity between specific bud-localizing mRNAs and non-specific unrelated RNAs was only moderate. The protruding helix E and the very C-terminus of She2p were shown to participate in *ASH1*-mRNA binding, and thus, provide RNA-binding specificity. Furthermore, bud-localizing mRNAs are bound by an extended surface region on She2p tetramers provided by the distinct RNA-binding motifs. Selective, but moderate reduction of *ASH1*-mRNA binding completely abolishes formation of a cytoplasmic transport complex and prevents mRNA localization *in vivo*. Thus, I conclude that She2p mediates the specificity for the interaction with *ASH1* mRNA.

However, it is absolutely clear that the complex of She2p and bud-localizing mRNAs must be stabilized during mRNP assembly *in vivo*. This assumption is supported by the observation that complexes consisting of only She2p and RNA are rather transient *in vitro* (this study). It is likely that the She2p:mRNA association is stabilized already in the nucleus. She2p is proposed to bind *ASH1* mRNA co-transcriptionally (Du et al., 2008), which might result in the formation of an unstable priming complex (Figure 36). Potential candidates for joining such a priming complex are the proteins Loc1p, Puf6p, and Khd1p. All of them are found in the nucleus, are reported to bind to *ASH1* mRNA, and are required for translational silencing of *ASH1* (Deng et al., 2008; Du et al., 2008; Gu et al., 2004; Irie et al., 2002; Long et al., 2001; Paquin et al., 2007; Shen et al., 2009). However, none of these proteins is known to bind to all four *ASH1*-localization elements. Thus, it remains unclear how the She2p:mRNA interaction might be stabilized in the nucleus. Preliminary studies indicate that the strictly nucleolar protein Loc1p interacts with all *ASH1* zipcodes (this study, data not shown). Loc1p in turn,

DISCUSSION

was recently identified to interact with She2p *in vivo* (Shen et al., 2009) and *in vitro* (this study, data not shown). Thus, one can assume that Loc1p might stabilize the interaction of She2p with its target mRNA, potentially by remodeling the mRNP complex. Puf6p binds to the *ASH1*-E3 zipcode as well as to She2p (Gu et al., 2004; Shen et al., 2009), whereas Khd1p binds to sequences adjacent to the *ASH1*-E1 zipcode (Irie et al., 2002). Their potential nuclear role in stabilization of She2p:mRNA complexes, however, remains unknown. Both proteins are also supposed to accompany the cytoplasmic translocation complex, thereby preventing premature *ASH1* translation (Gu et al., 2004; Paquin et al., 2007). By this means, Khd1p and Puf6p could positively affect the complex stability of She2p and the E1 and E3 zipcode, respectively (Figure 36).

In the cytoplasm, the adapter protein She3p is proposed to strengthen the She2p:mRNA complex through interaction of its C-terminal domain with She2p (Figure 36) (Böhl et al., 2000; Long et al., 2000). However, a stabilizing effect of She3p on mRNA binding by She2p remains to be demonstrated *in vitro*. Unfortunately, any attempt to purify stable She3p fragments for corresponding *in vitro* studies failed due to rapid protein degradation (data not shown). Likewise, short She3p peptides, which were apparently identified to interact with She2p by screening a She3p-peptide library, could not be bound to She2p (data not shown). Thus far, the She3p-interacting region of She2p is unknown. It was suggested previously that She3p interacts with a heterologous region consisting of the hydrophobic upper surface of She2p and mRNA (Gonsalvez et al., 2003; Niessing et al., 2004). This idea is based on the identification of a yeast strain encoding a She2p variant with an exchange of amino acid leucine 130 into serine (Gonsalvez et al., 2003). The She2p mutant failed to interact with She3p *in vivo*. Consistently with this, the She2p-L130Y mutant showed reduced *ASH1*-E3 binding (Niessing et al., 2004 and this study). However, with the identification of She2p-tetramer formation through the upper surface region, I conclude that the previously observed functional defects directly resulted from a disrupted oligomerization state of She2p. Thus, the upper surface of She2p is extremely unlikely to interact directly with She3p.

Besides the upper surface, there is only one additional small hydrophobic patch remaining that could be involved in She3p binding. Interestingly, this region is located on the lower half of She2p and involves parts of the recently identified non-classical nuclear localization signal (Figure 4) (Shen et al., 2009). In principle, She2p could interact through a similar surface region both with the importin α Srp1p and She3p, since both associations would be restricted to different time points during mRNP assembly. In this scenario, cargo-free She2p is imported

DISCUSSION

into the nucleus via its interaction with Srp1p. In the nucleus, She2p selectively recognizes bud-localizing mRNAs, traverses the nucleolus, and is exported to the cytoplasm in an mRNA cargo-dependent manner. In the cytoplasm, the She2p:mRNA complex is stabilized by interaction with She3p through a hydrophobic patch at the lower half of She2p. However, it is also possible that She3p interacts with a different region in She2p. Likewise, She3p might recognize a heterologous surface of She2p and mRNA, which may be located not on the top of She2p, but rather on the side.

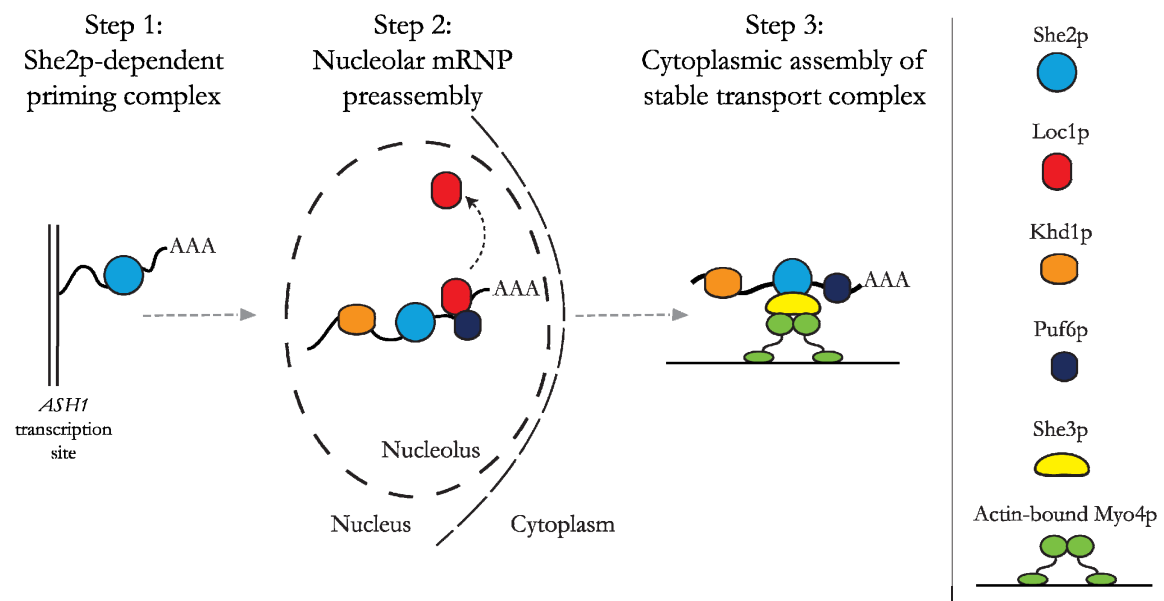


Figure 36: Sequential *ASH1*-mRNP assembly in *S. cerevisiae*. Step 1: She2p binds *ASH1* mRNA co-transcriptionally to form a priming complex. Step 2: The priming complex is passed through the nucleolus and is read-out by the *trans*-acting factors Loc1p, Khd1p, and Puf6p, which results in assembly of a translationally silent mRNP. Before the pre-mRNP is exported into the cytoplasm, Loc1p dissociates from the complex and remains in the nucleolus. Step 3: In the cytoplasm, the *ASH1* pre-mRNP associates with She3p, which stabilizes the complex. She3p itself is bound to Myo4p, which then transports the stable, translationally repressed mature mRNP to the bud tip.

The idea that an RNA-priming complex might form upon mRNA recognition might also apply to *Drosophila* Staufen. In a fundamental study, Staufen was reported to bind to RNA with extensive secondary structures, such as U1 and U2 RNAs, adenoviral VA1 RNA, and the 3' UTR of *bicoid* mRNA (St Johnston et al., 1992). RNAs without any predicted secondary structures were not bound. Interestingly, comparable attributes also apply to She2p (chapter 3.1). *In vivo* injection experiments in *Drosophila* embryos using a variety of RNAs showed that Staufen-containing mRNPs only form and localize efficiently upon injection of the *bicoid* 3' UTR, the target mRNA of Staufen (Ferrandon et al., 1994). Injection of highly concentrated heterologous stem-loop RNAs resulted in the formation of a few, yet immobile particles.

Based on this finding the authors concluded that the interaction of Staufén with *bicoid* is specific and that the interaction of Staufén with non-specific RNAs occurs in a fundamentally distinct way. Thus, it might be possible that also Staufén forms an initial and specific priming complex, which is interpreted and stabilized by accessory factors during mRNP assembly.

3.5 She2p binds RNA in multiple modes

Members of the Pumilio/FBF (PUF) family of RNA-binding proteins are known to recognize distinct variations of the PUF consensus sequence in the 3' UTR of their target mRNAs (Wickens et al., 2002). DNA-microarray analyses of five PUF proteins from *S. cerevisiae* indicate that each of them binds to a distinct and highly specific subset of 40 to 220 mRNAs (Gerber et al., 2004). Thus, it can be assumed that the respective PUF proteins bind to its different target mRNAs with high specificity. Analysis of the binding specificity of yeast Puf4p revealed that it has an only two-fold higher affinity for its native RNA target than for a Puf3p-specific target site (Miller et al., 2008). Similarly, recombinant Puf4p bound with 25-fold higher affinity to its own target than to a Puf5p (Mpt5p) target (Hook et al., 2007). Puf5p, in turn, displayed an only 3.5-fold preference for its own site. These findings indicate that the observed high specificity of PUF proteins for target mRNAs *in vivo* does not directly result from a high-affinity interaction. It is rather suggested that PUF proteins achieve specificity for their target mRNAs by applying a distinct binding mode, which might be further strengthened by association of accessory factors *in vivo*. This idea is supported by the recent finding that a single PUF protein from *C. elegans* (PUF-11) recognizes three different classes of RNAs by employing distinct RNA-binding modes with different RNAs (see also chapter 3.1) (Koh et al., 2009). The individual binding modes of PUF-11 rely on the flexibility in the central region of the protein and base flipping of the RNA and are adjusted according to the nucleotide composition in the respective RNA target.

The above mentioned examples of the PUF-protein family demonstrate a mechanism how a single RNA-binding protein achieves specificity to a number of different RNA targets. Although PUF proteins recognize specific nucleotides in the target RNA, they are still able to bind to large set of RNAs and to discriminate between different classes of RNAs. A comparable mechanism of RNA binding might also apply to She2p for the specific binding of a number of bud-localizing target mRNAs *in vivo*. In analogy to PUF-11 (Koh et al., 2009), it is rather unlikely that She2p achieves high affinity to a number of targets by accomplishing

DISCUSSION

several well-defined atomic contacts to every single zipcode. Instead, less coordination on the atomic level would be required for selective target recognition on the basis of an extended RNA-binding surface with moderate affinity. In addition, secondary structural features in the respective target mRNA might support specific interaction with She2p. Also potential influence of a degenerate RNA motif, as proposed by Jambhekar et al., 2005 and Olivier et al., 2005, cannot be excluded, even though to a lesser extent. Such an RNA-recognition mode would not require high sequence conservation on the zipcode level, which is supported by the observation that there is only little similarity on the sequence level among yeast zipcodes (Jambhekar et al., 2005; Olivier et al., 2005). Instead, a potentially conserved tertiary structure of bud-localizing mRNAs might be required for correct mRNA localization. The priming complex of RNA-bound She2p may then be stabilized by accessory factors, like cytoplasmic She3p, if the overall topology indicates specific mRNA binding. In this case, assembly with She3p triggers the formation of a functional mRNP.

In a recent genome-wide study the RNA targets of selected 40 RNA-binding proteins (out of more than 500) in yeast were systematically analyzed (Hogan et al., 2008). They found that at least 33 of these 40 proteins, including She2p, associate with specific sets of a few to several hundred mRNAs each. Thus, specific mRNA recognition through multiple binding modes might be a more general mechanism that could account for the assembly of other mRNPs as well.

4 Materials and Methods

4.1 Consumables and chemicals

Unless specified otherwise, all common chemicals were obtained in the highest available quality from Roth (Karlsruhe, Germany), Merck (Darmstadt, Germany), or Sigma (Deisenhofen, Germany). Enzymes and nucleotides for molecular biology and molecular weight markers for electrophoretic analyses were purchased from Fermentas (St. Leon-Rot, Germany), New England Biolabs (Frankfurt, Germany), Invitrogen (Karlsruhe, Germany), or Sigma (Deisenhofen, Germany). Reagents to prepare media for bacterial and yeast cultivation were obtained from Becton Dickinson (Heidelberg, Germany) and Sigma (Deisenhofen, Germany). Chromatography equipment, columns and media were purchased from GE Healthcare (Munich, Germany). Transfer membranes for western blotting and filter-binding experiments were obtained from Roth (Karlsruhe, Germany). Radioactive nucleotides (γ - ^{32}P -ATP and 5'- ^{32}P -pCp) were obtained from GE Healthcare (Munich, Germany) and Hartmann Analytik (Braunschweig, Germany). Crystallization screens, crystallization grade reagents, and crystallization tools were purchased from Hampton Research (Aliso Viejo, USA), Qiagen (Hilden, Germany), and Jena Bioscience (Jena, Germany). DNA oligonucleotides were ordered from Thermo Fisher (Ulm, Germany). RNA oligonucleotides for crystallization and biochemical assays were synthesized by Thermo Fisher Scientific/Dharmacon (USA).

4.2 Oligonucleotides

4.2.1 DNA oligonucleotides

#	oligonucleotide	sequence (5'-3')
S2-1	She2p_Crev_XhoI	TTACTCGAGTCAGTTTTTCAATTTACCAAATTTG
S2-2	She2p_243rev_Stop_XhoI	TAACTCGAGTTATTTACCAAATTTGTCATGCC
S2-3	She2_Nfor_BamHI	AAAGGATCCATGAGCAAAGACAAAGATATC
S2-4	She2_aa173_rev	TCTCCGATATTCGTTCCGTCAACTTCTAAATCTAGCAGGTTGCTGC
S2-5	She2_aa184_for	TGCTAGATTTAGAAGTTGACGGAACGAATATCGGAGAAACTG
S2-10	She2_DloopBC_80_rev	TACCTTGTCAAACGAATCTTCATTGATAAACTCTGATGCATT
S2-11	She2_DloopBC_90_for	TCAGAGTTTATCAATGAAGATTCGTTTGACAAGGTAATTTTACC
S2-12	She2_S120Y_for	CTGAACTATTATCTAACGCAGTactTACAAAAGGAAATTTTATCTA AAACTTTGAACG
S2-13	She2_S120Y_rev	CGTTCAAAGTTTTAGATAAAAATTCCTTTTGTAAGtACTGCGTTAG ATAATAGTTCAG

MATERIALS AND METHODS

#	oligonucleotide	sequence (5'-3')
S2-17	She2_nt218_for_NsiI	AAAAAATGCATCAGAGTTTATCAATGAAGG
S2-18	She2_Crev_StuI	AAAAAAGGCCTCAGTTTTTCAATTTACC
S2-19	She2_Crev_XbaI	AAATCTAGAGTTTTTCAATTTACCAAATTTGTCATGC
S2-20	She2_240rev_XbaI	AAATCTAGATTTGTCATGCCATTTAGTAGCGAC
S2-21	She2_240rev_StuI	AAAAAAGGCCTCATTTGTCATGCCATTTAGTAGCGAC
S2-22	She2_240rev_XhoI	TAACTCGAGTTTGTCTATGCCATTTAGTAGCGAC
S2-23	She2_L130Y_for	GGAAATTTTATCTAAACTTACAACGAGGACCTAACGC
S2-24	She2_L130Y_rev	GCGTTAGGTCCCTCGTTGTAAGTTTTAGATAAAATTTCC
S2-26	XbaI_She2DHelix_3myc_r	GTCTCTAGACAATTTACCAAATTTGTCATGCC
S2-27	She2_6for_BamHI	AAAGGATCCATGGATATCAAAGTCACTCCTGG

4.2.2 RNA oligonucleotides

RNA	Base no. relative to start codon	Sequence (5'-3')	Produced by
<i>ASH1</i> E1	623-701	AAGACUAUGUAAAAUACGCGAAGAAGUGGCU CAUUUCAAGCCAUUAAGUAUACCCAACUUAAC UAAUAAUCAAUAUA	<i>in vitro</i> transcription (this study)
<i>ASH1</i> E2A	1109-1185	UUGCGAAUAGAGACAUUCUAUCGAACAAUUC AAAUCUAAUGUAAGGAAACCAUCUAAGAACAA AAUCUCAAGCAA	<i>in vitro</i> transcription (this study)
<i>ASH1</i> E2B domain 1	1263-1320	AUCUUAUCUCCAUCUCCUCCACACCGACGA AAAGUGGCAAGAUGAGAUAAGAUA	<i>in vitro</i> transcription (this study)
<i>ASH1</i> E3	1750-1867	GAGACAGUAGAGAAUUGAUACAUGGAUAACUG AAUCUCUUUAACUAAUAAGAGACAUAUACAC GAAACAAUUGUACAUAUUCUCCUUGUCUGUG CUAAUAAACUACAAUAAAAA	<i>in vitro</i> transcription (this study)
<i>EAR1</i>	1572-1621	CGAAGAUGAAAAUGAUCACGAGAGAGAUCCCG AACAAUUUUCAGAAUUUG	chemical synthesis (Dharmacon); 5' biotinylated
<i>WSC2N</i>	430-465	AGUUCAAAAACGUCCACGAAAUUGGACACGAA AACU	chemical synthesis (Dharmacon); 5' biotinylated
<i>ASH1</i> E1min	635-683	AAUACGCGAAGAAGUGGCUCAUUUCAAGCCA UUAAGUAUACCCAACUU	chemical synthesis (Dharmacon)
<i>ASH1</i> E2Bmin (36 bases)	1279-1314	CCCUCCACACCGACGAAAAGUGGCAAGAUGAG AUCA	chemical synthesis (Dharmacon)
<i>ASH1</i> E2Bmin (21 bases)	G+1283- 1303	GCCACACCGACGAAAAGUGGC	chemical synthesis (Dharmacon)
E2Bshort-01		GGCCACACCGACGAAAAGUGGCC	chemical synthesis (Dharmacon)
E2Bshort-02		AGGCCACACCGACGAAAAGUGGCC	chemical synthesis (Dharmacon)
E2Bshort-03		GGCCACACCGACGAAAAGUGGCCA	chemical synthesis (Dharmacon)
E2Bshort-04		UCCACACCGACGAAAAGUGGA	chemical synthesis (Dharmacon)

MATERIALS AND METHODS

RNA	Base no. relative to start codon	Sequence (5'-3')	Produced by
E2Bshort-05		CCACACCGACGAAAAGUGG	chemical synthesis (Dharmacon)
E2Bshort-06		CCACACCGACGAAAAGUGGC	chemical synthesis (Dharmacon)
E2Bshort-07		ACCACACCGACGAAAAGUGGU	chemical synthesis (Dharmacon)
E2Bshort-08		GCACACCGACGAAAAGUGC	chemical synthesis (Dharmacon)
E2Bshort-09		ACCCACCGACGAAAAGGGU	chemical synthesis (Dharmacon)
HIV-1 TAR (16 bases)		CGAGCCUGGGAGCUCG	chemical synthesis (Dharmacon)
HIV-1 TAR (57 bases)		GGUCUCUCUGGUUAGACCAGAUCUGAGCCUGG GAGCUCUCUGGCUAACUAGGGAACC	<i>in vitro</i> transcription (this study)
U1snRNA hairpin II		UUAUCCAUUGCACUCCGAUGU	chemical synthesis (Dharmacon)
MS2 stem loop		ACAUGAGGAUUACCCAUGU	chemical synthesis (Dharmacon)
Poly(A) ₂₀		AAAAAAAAAAAAAAAAAAAAA	chemical synthesis (Dharmacon)
9-mer motif		CGACGAAAA	chemical synthesis (Dharmacon); 5' biotinylated for SPR analysis
7-mer motif		ACGAAAA	chemical synthesis (Dharmacon)
6-mer motif		CGACGA	chemical synthesis (Dharmacon); 5' biotinylated for SPR analysis
3-mer motif		CGA	chemical synthesis (Dharmacon); 5' biotinylated for SPR analysis
TAR+6-mer motif (sense)		CGAGCCACUCCGACGAGCUCG	chemical synthesis (Dharmacon)
TAR+6-mer motif (antisense)		CGAGCCACUCGUCGUCGUCG	chemical synthesis (Dharmacon)

4.2.3 DNA oligonucleotides to generate templates for *in vitro* transcription

dsDNA templates for *in vitro*-transcription reactions were generated by PCR using the following oligonucleotides. Underlined sequences correspond to the T7 promoter-consensus sequence that was added to the 5' end of the coding strand through the respective forward primer.

MATERIALS AND METHODS

#	oligonucleotide	sequence (5'-3')
Tx-1	ASH1_E1_for	<u>TAATACGACTCACTATAGG</u> GAGAAGACTATGTTAAAATACGCGAAG
Tx-2	ASH1_E1_rev	TTATTTTGATTATTAGTTAAGTTG
Tx-3	ASH1_E2A_for	<u>TAATACGACTCACTATAGG</u> GAGTTGCGAATAGAGACATTCTATCG
Tx-4	ASH1_E2A_rev	AGTTGCTTTGAGATTTTGTTCCTAC
Tx-10	ASH1_E3_118nt_for	<u>TAATACGACTCACTATAGG</u> GAGACAGTAGAGAATTGATACATGG
Tx-11	ASH1_E3_118nt_rev	TTTTTATTTGTAGTTTATTTAGC
Tx-12	ASH1_E2B_D1_for	<u>TAATACGACTCACTATAGG</u> GAGATCTTCATCTCCATCTCCCTCCAC
Tx-13	ASH1_E2B_D1_rev	TGATCTTGATCTCATCTTGCC
Tx-28	TAR-57_up	<u>TAATACGACTCACTATAGG</u> GGTCTCTCTGGTTAGACCAGATCTGAG CCTGGGAGCTCTCTGGCTAACTAGGGAACC
Tx-29	TAR-57_down	GGTTCCCTAGTTAGCCAGAGAGCTCCAGGCTCAGATCTGGTCTAA CCAGAGAGACCCCTATAGTGAGTCGTATTA

4.3 Plasmids

4.3.1 Plasmids for protein expression in *E. coli*

#	insert	vector	restriction sites	primer	tag	comments
p01	She2p	pGEX-6P-1	BamH/XhoI	S2-3 + S2-1	N-GST	full length (246 aa); (Niessing et al., 2004)
p02	She2p-ΔCys	pGEX-6P-1	BamH/XhoI	S2-3 + S2-1	N-GST	C14S, C68S, C106S, C180S; created by Dierk Niessing
p03	She2p-(6-246)-ΔCys	pGEX-6P-1	BamH/XhoI	S2-27 + S2-1	N-GST	Δaa1-5; C14S, C68S, C106S, C180S
p04	She2p-(6-243)-ΔCys	pGEX-6P-1	BamH/XhoI	S2-27 + S2-2	N-GST	Δaa1-5; Δaa244-246; C14S, C68S, C106S, C180S
p05	She2p-(6-239)-ΔCys	pGEX-6P-1	BamH/XhoI	S2-27 + S2-22	N-GST	Δaa1-5; Δaa240-246; C14S, C68S, C106S, C180S
p06	She2p-(ΔC-term)	pGEX-6P-1	BamH/XhoI	S2-3 + S2-22	N-GST	Δaa241-246
p08	She2p-L130Y	pGEX-6P-1	BamH/XhoI	S2-3 + S2-1	N-GST	site-directed mutagenesis using oligos S2-23 + S2-24
p09	She2p-S120Y	pGEX-6P-1	BamH/XhoI	S2-3 + S2-1	N-GST	site-directed mutagenesis using oligos S2-12 + S2-13
p10	She2p-(N36S,R63K)	pGEX-6P-1	BamH/XhoI	-	N-GST	(Du et al., 2008)
p11	She2p-ΔhelixE	pGEX-6P-1	BamH/XhoI	S2-3 + S2-1	N-GST	Δaa174-183; deleted using oligos S2-4 + S2-5
p13	She2p-Δloop	pGEX-6P-1	BamH/XhoI	S2-3 + S2-1	N-GST	Δaa81-89; internal deletion created using oligos S2-10 + S2-11

4.3.2 Yeast plasmids

#	insert	vector	restriction sites	template	primer	tag	comments
RJP132	<i>ASH1</i>	YEplac195					URA marker; (Long et al., 1997)
RJP138	-	YCplac22					TRP marker; (Gietz and Sugino, 1988)
RJP145	-	YCplac111					LEU marker; (Gietz and Sugino, 1988)
RJP428	<i>SHE2-3myc</i>	pUC9				C-3xmyc	(Jansen et al., 1996)
p15	<i>SHE2-3myc</i>	YCplac22	HindIII	RJP428		C-3xmyc	subcloned from RJP428 via HindIII
p16	<i>SHE2-L130Y-3myc</i>	YCplac22	NsiI/XbaI	p08	S2-17 + S2-19	C-3xmyc	She2-L130Y-PCR fragment cloned into p15
p17	<i>SHE2-ΔhelixE-3myc</i>	YCplac22	NsiI/XbaI	p11	S2-17 + S2-26	C-3xmyc	SHE2-ΔhelixE-PCR fragment cloned into p15
p18	<i>SHE2-Δ(C-term)-3myc</i>	YCplac22	NsiI/XbaI	p06	S2-17 + S2-20	C-3xmyc	SHE2-Δ(C-term)-PCR fragment cloned into p15
RJP916	<i>SHE2</i>	YCplac111					(Du et al., 2008)
p19	<i>She2-L130Y</i>	YCplac111	NsiI/StuI	p08	S2-17 + S2-18	-	She2-L130Y-PCR fragment cloned into RJP916
p20	<i>SHE2-ΔhelixE</i>	YCplac111	NsiI/StuI	p11	S2-17 + S2-18	-	SHE2-ΔhelixE-PCR fragment cloned into RJP916
p21	<i>SHE2-Δ(C-term)</i>	YCplac111	NsiI/StuI	p06	S2-17 + S2-21	-	SHE2-Δ(C-term)-PCR fragment cloned into RJP916

4.4 Bacterial strains

Strain	Essential genotype	Source
XL-1 blue	<i>recA endA1 gyrA96 thi-1 hsdR17 supE44 relA1 lac</i> [F' <i>proAB lacI^q ZΔM15 Tn10 (Tet^r)</i>]	Stratagene (La Jolla, USA)
BL21 Star (DE3)	F- <i>ompT hsdSB (rB-mB-) gal dcm rne131</i> (DE3)	Invitrogen (Karlsruhe, Germany)

4.5 Yeast strains

Strain	Essential genotype	Source
RJY 2053	<i>MAT alpha; his3Δ1; leu2Δ0, lys2Δ0, ura3Δ0; she2::kanMX4</i>	EUROSCARF (Frankfurt, Germany)
RJY 3364	<i>MAT a, his3, leu2, ade2, trp1, ura3, HO-ADE2, HO-CAN1, HA3-SHE3, she2::natNT2</i>	Gift from R.-P. Jansen (Tübingen)

MATERIALS AND METHODS

Strain	Essential genotype	Source
y01	<i>MAT alpha; his3Δ1;leu2Δ0, lys2Δ0, ura3Δ0; she2::kanMX4, YEplac195-ASH1</i> (RJP132), <i>YCplac111</i> (RJP145)	this study
y02	<i>MAT alpha; his3Δ1;leu2Δ0, lys2Δ0, ura3Δ0; she2::kanMX4, YEplac195-ASH1</i> (RJP132), <i>YCplac111-SHE2</i> (RJP916)	this study
y03	<i>MAT alpha; his3Δ1;leu2Δ0, lys2Δ0, ura3Δ0; she2::kanMX4, YEplac195-ASH1</i> (RJP132), <i>YCplac111-SHE2-ΔhelixE</i> (p20)	this study
y04	<i>MAT alpha; his3Δ1;leu2Δ0, lys2Δ0, ura3Δ0; she2::kanMX4, YEplac195-ASH1</i> (RJP132), <i>YCplac111-SHE2-ΔC-term</i> (p21)	this study
y05	<i>MAT alpha; his3Δ1;leu2Δ0, lys2Δ0, ura3Δ0; she2::kanMX4, YEplac195-ASH1</i> (RJP132), <i>YCplac111-SHE2-L130Y</i> (p19)	this study
y06	<i>MAT a, his3, leu2, ade2, trp1, ura3, HO-ADE2, HO-CAN1, HA3-SHE3, she2::natNT2, YCplac22</i> (RJP138)	this study
y07	<i>MAT a, his3, leu2, ade2, trp1, ura3, HO-ADE2, HO-CAN1, HA3-SHE3, she2::natNT2, YCplac22-SHE2-myc3</i> (p15)	this study
y08	<i>MAT a, his3, leu2, ade2, trp1, ura3, HO-ADE2, HO-CAN1, HA3-SHE3, she2::natNT2, YCplac22-SHE2-ΔhelixE-myc3</i> (p17)	this study
y09	<i>MAT a, his3, leu2, ade2, trp1, ura3, HO-ADE2, HO-CAN1, HA3-SHE3, she2::natNT2, YCplac22-SHE2-ΔC-term-myc3</i> (p18)	this study
y10	<i>MAT a, his3, leu2, ade2, trp1, ura3, HO-ADE2, HO-CAN1, HA3-SHE3, she2::natNT2, YCplac22-SHE2-L130Y-myc3</i> (p16)	this study

4.6 Media and supplements

Medium	Composition	Purpose
LB medium	1 % (w/v) bacto tryptone, 0.5 % (w/v) bacto yeast extract, 0.5 % (w/v) NaCl (+1.5 % (w/v) agar for plates)	Growth medium for <i>E. coli</i>
Medium A (per liter)	100 ml M9 medium (10x), 10 ml trace elements solution (100x), 20 ml 20 % (w/v) glucose, 1 ml 1 M MgSO ₄ , 0.3 ml 1 M CaCl ₂ , 1 ml biotin (1 mg/ml), 1 ml thiamin (1 mg/ml), appropriate antibiotic(s)	Growth medium for <i>E. coli</i> ; ¹⁵ N-labeling of proteins for NMR experiments
YPD medium	1 % (w/v) yeast extract, 2 % (w/v) bacto peptone, 2 % (w/v) glucose (+2 % (w/v) agar for plates)	Growth medium for <i>S. cerevisiae</i>
Synthetic Complete (SC) medium	0.67 % (w/v) yeast nitrogen base (w/o amino acids), 2 % (w/v) glucose, 0.06 % (w/v) CSM dropout mix (includes all essential amino acids and nucleobases except the amino acids or nucleobases used as auxotrophy markers, i.e. leucine, tryptophane, histidine, adenine, and uracil)	Selective growth medium for <i>S. cerevisiae</i>

Supplement	Stock solution	Final concentration
M9 medium (10x) (per liter)	60 g Na ₂ HPO ₄ , 30 g NaH ₂ PO ₄ , 5 g NaCl, 5 g ¹⁵ NH ₄ Cl	1x (Supplement for medium A)
Trace elements solution (100x) (per liter)	5 g EDTA, 0.83 g FeCl ₃ x 6 H ₂ O, 84 mg ZnCl ₂ , 13 mg CuCl ₂ x 2 H ₂ O, 10 mg CoCl ₂ x 6 H ₂ O, 10 mg H ₃ BO ₃ , 1.6 mg MnCl ₂ x 6 H ₂ O	1x (Supplement for medium A)
Adenine	7.5 g/l	75 mg/l

Supplement	Stock solution	Final concentration
Uracil	2 g/l	20 mg/l
Histidine	2 g/l	20 mg/l
Leucine	10 g/l	100 mg/l
Tryptophane	5 g/l	50 mg/l
Ampicillin	100 mg/ml	100 mg/l
Kanamycin	50 mg/ml	50 mg/l
IPTG	1 M	0.3 mM

4.7 Antibodies

Primary antibody	Source	Dilution / Application	Supplier
anti-HA (3F10)	rat	1:1000 / Western Blot	Roche Applied Science (Mannheim, Germany)
anti-myc (9E10)	mouse	1:1000 / Western Blot	Roche Applied Science (Mannheim, Germany)
anti-myc (9E11)	mouse	100 µg antibody coupled to 250 µl protein-G beads / Co-IP	Acris Antibodies (Herford, Germany)
anti-She2p (323/4)	rabbit	1:1000 / IF	(Du et al., 2008)
anti-She2p-1C3	rat	1:50 (hybridoma supernatant) / Western Blot	Generated in collaboration with E. Kremmer (Helmholtz Zentrum München, Germany)

Secondary antibody	Source	Dilution / Application	Supplier
Alexa®488 anti-rabbit-IgG	goat	1:250 / IF	MoBiTec (Göttingen, Germany)
anti-mouse-IgG-HRPO	goat	1:3000 / Western	BioRad (Munich, Germany)
anti-rat-IgG-HRPO	rabbit	1:10000 / Western	Abcam (Cambridge, USA)

4.8 Bioinformatic tools and software

Physical and chemical parameters (molecular weight, theoretical isoelectric point, extinction coefficient) of recombinant proteins were calculated using the ProtParam tool from the ExPASy Proteomics Server (<http://www.expasy.org/>) (Gasteiger et al., 2003).

Secondary structures of RNA were predicted using the MC-Fold algorithm (<http://www.major.irc.ca/MC-Fold/>) (Parisien and Major, 2008).

RNA-filter binding experiments were analyzed and quantified using the ImageQuant software from GE Healthcare and the software KaleidaGraph from Synergy Software (Reading, USA). Co-immunoprecipitation experiments were quantified using the analysis software MultiGauge from FUJIFILM.

SAXS data were analyzed and processed with different programs from the ATSAS 2.1 software package (Konarev et al., 2006). *Ab initio* modeling was performed using GASBOR

(Svergun et al., 2001). Averaging of multiple GASBOR models and superposition of the She2p-crystal structure with SAXS models was performed with SUPCOMP and DAMAVER (Volkov and Svergun, 2003).

All figures of protein structures were generated with the program PyMOL from DeLano Scientific (Palo Alto, USA) (DeLano, 2004).

4.9 Cloning, expression and purification methods

4.9.1 Cloning

Standard methods for molecular biology were performed as described (Sambrook and Russell, 2001). This includes polymerase chain reaction for amplification of DNA fragments, restriction digest of DNA, ligation of DNA fragments, generation of chemically competent *Escherichia coli* cells, transformation of plasmid DNA, isolation of plasmid DNA from *E. coli* cells, and analysis of DNA by agarose gelelectrophoresis. Generally, the isolation of plasmid DNA and the purification of DNA fragments was performed with the NucleoSpin-Plasmid and the NucleoSpin-Extract II kits from Macherey-Nagel (Düren, Germany), respectively. Likewise, restriction enzymes were used according to the manufacturer's instruction. Point mutations and deletion of short fragments within a gene were generated by PCR-based site-directed mutagenesis (Ho et al., 1989). DNA sequencing was performed by Eurofins Medigenomix (Martinsried, Germany). The cloning strategies for all plasmids used for either recombinant protein expression in *E. coli* or functional analysis in *S. cerevisiae* are summarized in chapter 4.3. A list of all oligonucleotides that were used for cloning is shown in chapter 4.2.1.

4.9.2 Recombinant protein expression in *E. coli*

All variants of GST-She2-fusion proteins used in this study were recombinantly expressed in *E. coli* BL21 Star (DE3) cells. Cells were grown at 37 °C in LB medium with the appropriate antibiotics to an OD₆₀₀ of 0.4-0.6 and then cooled down to 18 °C. Protein synthesis was induced by adding 0.3 mM IPTG and carried out at 18 °C for 12-16 hours. The cells were harvested by centrifugation at 4 °C, resuspended in lysis buffer (for all buffers, see chapter 4.9.4), flash frozen in liquid nitrogen, and stored at -80 °C.

4.9.3 Isotopic ^{15}N -labeling for NMR spectroscopy

For exploratory NMR spectroscopy experiments, GST-She2p was ^{15}N -labeled during protein expression in minimal medium containing $^{15}\text{NH}_4\text{Cl}$. For the starter culture, 20 ml Medium A (see chapter 4.6) containing ampicillin was inoculated with a single colony of freshly transformed *E. coli* BL21 Star (DE3) cells carrying the plasmid p01 (chapter 4.3.1) and grown over night at 37 °C. With the 20 ml over-night culture, one liter of Medium A containing ampicillin was inoculated and grown at 37 °C to an OD_{600} of 0.8. After cooling the culture on ice, over-expression of GST-She2p was induced by adding 0.3 mM IPTG. The cells were grown for additional 16 hours at 18 °C and then harvested by centrifugation at 4 °C. The pellet was resuspended in lysis buffer (for all buffers, see chapter 4.9.4) and flash frozen in liquid nitrogen before storage at -80 °C.

4.9.4 Purification of recombinant She2p variants

All She2p variants used in this study (see chapter 4.3.1) were expressed as GST-fusion proteins. After harvesting of the cells and resuspending the pellet in lysis buffer (see table below), the cells were disrupted by sonication. Cell debris was removed by centrifugation at 4 °C (SS34 rotor, 15000 rpm, 1 h). For affinity purification, the supernatant was applied onto an equilibrated GSTrap 5 ml column. Subsequently, the column was extensively washed with the GST-wash buffers 1 and 2. The washing step with 1 M NaCl-containing buffer proved to be essential for eliminating potential She2p-associated nucleic acids. The GST-fusion protein was eluted with GST-elution buffer. To remove the GST tag from She2p, 50 µg PreScission Protease (GE Healthcare, Munich) was added to the elution fraction, which was then dialyzed against dialysis buffer for 12 h at 4 °C. Afterwards, the cleaved tag was removed from the solution by applying the dialysis fraction onto a GSTrap column. The flow through fraction, which contains the untagged She2 protein, was loaded onto a HiTrap-Q column for anion exchange chromatography. Prior to the experiment, the ion exchange column was pre-equilibrated with buffer A. The protein was eluted with a gradient of 20 column volumes ranging from buffer A to buffer B. Peak fractions were analyzed by SDS-PAGE (Laemmli, 1970), pooled and concentrated using centrifugal filter devices (Amicon Ultra, Millipore, Billerica, US). For the final purification step, the concentrated protein solution was loaded onto a Superdex S200 16/60 size exclusion chromatography column equilibrated with size-exclusion buffer. Different size-exclusion buffers were chosen dependent on the following experiment. After SDS-PAGE analysis, the peak fractions were concentrated and either subsequently used or flash frozen and stored at -80 °C. For crystallization experiments, She2p

was purified always directly prior the experiment, whereas for later biochemical analyses the protein was stored in small aliquots at -80 °C for a maximum duration of three months. For She2p:RNA co-crystallization experiments or for RNA-binding experiments, the protein samples should lack any RNase contaminants. To ensure this, the buffer for the final size-exclusion chromatography step was prepared with RNase-free chemicals and RNase-free water (DEPC treated). In case of cysteine-free mutant proteins, DTT was omitted from the respective buffers.

Buffer for She2p purification	Composition
lysis buffer	20 mM HEPES, 1 M NaCl, 2 mM DTT, 1x Complete EDTA-free protease-inhibitor cocktail (Roche Applied Science), pH 7.4 (8 °C)
GST-wash buffer 1	20 mM HEPES, 0.5 M NaCl, 2 mM DTT, pH 7.4 (8 °C)
GST-wash buffer 2	20 mM HEPES, 1 M NaCl, 2 mM DTT, pH 7.4 (8 °C)
GST-elution buffer	20 mM HEPES, 0.5 M NaCl, 2 mM DTT, 30 mM glutathione, pH 7.4 (8 °C)
dialysis buffer	20 mM HEPES, 150 mM NaCl, 2 mM DTT, pH 7.9 (8 °C)
buffer A	10 mM HEPES, 150 mM NaCl, 2 mM DTT, pH 7.9 (8 °C)
buffer B	10 mM HEPES, 1 M NaCl, 2 mM DTT, pH 7.9 (8 °C)
size-exclusion buffer 1 (crystallization and RNA-binding)	10 mM HEPES, 100 mM NaCl, 2 mM DTT, pH 7.4 (8 °C)
size-exclusion buffer 2 (AUC)	20 mM Tris, 150 mM NaCl, 5 mM DTT, 1 mM EDTA, pH 7.4 (8 °C)
size-exclusion buffer 3 (analytical SEC)	20 mM Tris, 150 mM NaCl, 5 mM DTT, pH 7.4 (8 °C)
size-exclusion buffer 4 (SAXS)	50 mM Na ₂ HPO ₄ /NaH ₂ PO ₄ , 100 mM NaCl, 2 mM DTT, pH 7.4 (8 °C)
size-exclusion buffer 5 (NMR)	20 mM Na ₂ HPO ₄ /NaH ₂ PO ₄ , 100 mM NaCl, 2 mM MgCl ₂ , pH 6.8 (8 °C)

4.10 Methods for protein analysis

4.10.1 Western blotting

Following protein separation by SDS-PAGE (Laemmli, 1970), proteins were transferred onto a PVDF membrane for 2 h at 60 V using a wet-blotting apparatus (BioRad, Munich, Germany). After transfer, the membrane was blocked with 10 % (w/v) milk powder in TBS-T (10 mM Tris pH 8.0, 150 mM NaCl, 0.05 % Tween 20) and subsequently incubated with the primary antibody (diluted in 1x TBS-T containing 1 % milk powder) over night at 4 °C. After three washing steps with TBS-T, the membrane was incubated with the secondary antibody (diluted in 1x TBS-T containing 1 % milk powder) for 1 h at RT. Excess of antibody was removed by 3 washing steps with TBS-T. Immuno-stained proteins were visualized using an ECL kit (Pierce by Thermo Scientific) followed by exposure of the membrane to light-sensitive films (GE Healthcare) and developing with a Kodak Xomat M35 developing

machine. Alternatively for the purpose of quantification, ECL-treated membranes were directly analyzed with the LAS-3000 mini imaging system (FUJIFILM).

4.10.2 Dynamic Light Scattering

Dynamic Light Scattering (DLS) can be used to determine the size distribution of small particles in solution. DLS functions by measuring the speed at which particles move under Brownian motion. In a DLS system, a laser passes through a measuring cell, which contains the protein sample, and the intensity of light scattered by the sample is measured. Since molecules in solution are constantly moving, the scattering intensity will vary. Using the autocorrelation function, the hydrodynamic radius of the particles can be determined. Thus, DLS can be used to directly evaluate the effect of the solvent on the behavior, e.g. the solubility, of a given molecule. This technique was successfully used to screen for appropriate buffer conditions for SAXS experiments. Before the measurement, purified She2p was dialyzed over night against various buffers containing different buffer substances, salt concentrations, or additives like glycerol and detergents. Immediately before the experiment, the protein samples were centrifuged at 20000x g for 30 minutes to separate insoluble particles. The measurements were performed with a 70 μ l sample volume-quartz cuvette at 20 °C in a 802 DLS spectrometer (Viskotec Corp., Houston, USA). By testing a series of protein concentrations ranging from 1.7 mg/ml to 14.4 mg/ml, potential concentration-dependent oligomerization effects or aggregation of the sample could be measured.

4.10.3 Analytical ultracentrifugation

Analytical ultracentrifugation (AUC) experiments with She2p-wild type and She2p-L130Y were performed in close collaboration with Dr. Klaus Richter from the department of chemistry at the Technical University Munich. Sample preparation was done in our laboratory, whereas the AUC runs and data evaluation was performed by Klaus Richter. AUC was performed using a Beckman XL-A instrument (Beckman Coulter, Krefeld, Germany) equipped with UV/VIS optical systems. For all experiments, proteins were measured in size-exclusion buffer 2 (see chapter 4.9.4) and had been pre-treated with 20 μ g/ml RNaseA (Fermentas, St. Leon-Rot, Germany). In order to investigate concentration-dependent oligomerization events, the following protein concentrations were used in sedimentation-equilibrium experiments: wild type-She2p: 1.7 μ M, 2.6 μ M, 3.5 μ M, 5.2 μ M, 7.3 μ M, 10.5 μ M, 15.0 μ M, 21.6 μ M, 30.7 μ M, and She2p (L130Y): 1.9 μ M, 2.6 μ M, 3.7 μ M, 5.2 μ M, 7.7 μ M, 10.8 μ M, 15.7 μ M, 22.3 μ M, 32.0 μ M. 6-channel centerpieces were filled with 105 μ l of sample volume and 15 μ l heptacosafuorotributylamine. After 48 hours at 4°C, equilibrium conditions

were obtained for wild type-She2p (9000 rpm) and She2p -L130Y (12500, 18000, 23000 rpm), as judged from stable sedimentation profiles. Data were analyzed with the ULTRASCAN 9.8 software package (Demeler, 2005). Sedimentation-velocity experiments were performed at 50000 rpm and 20 °C with two channel centerpieces (Ti-60 rotor), and data were acquired at 280 nm wavelength. The following protein concentrations were measured: She2p-wild type: 16.0 μ M; She2p-L130Y: 4.9 μ M, 12.9 μ M, 32.0 μ M. For analysis of sedimentation traces, the dc/dt approach in the ULTRASCAN software package was used. In order to accurately determine S-values, resulting dc/dt curves were fitted in SigmaPlot (Systat Software, San Jose, USA) to a Gaussian function. Despite the potentially complex association model, this approach still resulted in consistent values with little non-stochastic residuals remaining after the data fitting.

4.10.4 Analytical size-exclusion chromatography

Analytical size-exclusion chromatography experiments were performed with a Superose 12 10/300 GL column (GE Healthcare), which was calibrated for molecular weight calculation using the Molecular Weight Marker Kit (Sigma, Deisenhofen, Germany). For molecular weight determination, 0.5 ml protein solution of She2p-wild type and She2p-L130Y, respectively, at various protein concentrations was applied onto the column with a flow rate of 0.5 ml/min and analyzed in size-exclusion buffer 3 (see chapter 4.9.4). The absorbance of eluted protein was measured at 280 nm and peak fractions were analyzed for presence of She2p by SDS PAGE or western blotting using an α -She2p-1C3 antibody (0.2 μ M samples). She2p-containing peak fractions were integrated over the peak volume to calculate the real protein concentration on the column.

4.10.5 Monoclonal antibodies

In collaboration with Dr. Elisabeth Kremmer (Helmholtz Zentrum München), monoclonal antibodies from rat against GST-tagged She2p were generated using standard procedures. Antibodies were tested in an ELISA using recombinantly expressed GST-She2p or GST protein alone. Monoclonal antibodies reacting exclusively with GST-She2p were further analyzed in western blotting. Clones She2p-1C3, She2p-4G8 and She2p-5A1 (rat IgG2a) recognize wild-type and mutant She2p specifically. In this work, She2p-1C3 was used for detection of She2p variants in western blotting.

4.11 Biochemical analysis of She2p:RNA interactions

4.11.1 *In vitro* transcription and RNA purification

For working with RNA, only RNase-free water and RNase-free chemicals were used. RNase-free water was prepared by treatment with the RNase inhibitor diethylpyrocarbonate (DEPC). In sterile bottles, DEPC was added to Millipore-H₂O to a final concentration of 0.05 % (v/v), incubated over night in a 40 °C water bath and then autoclaved for 20 minutes at 121 °C.

Chemically synthesized RNAs obtained from Thermo Fisher Scientific/Dharmacon (USA) (see chapter 4.2.2) were delivered in the stable 2'-ACE-protected form to prevent RNA hydrolysis. Deprotection of these oligonucleotides was performed according to the provided instructions. Then, the RNA was dissolved in RNase-free water and stored at -20 °C.

ASH1-zipcode RNA (E1, E2A, E2B, and E3; see chapter 4.2.2) and the HIV-1 TAR RNA (57 bases) for biochemical studies were generated by *in vitro* transcription using the MEGAscript T7 Kit (Applied Biosystems/Ambion, Austin, USA). Double-stranded DNA templates for the respective transcription reactions were generated by PCR (*ASH1* zipcodes) or by hybridization of complementary DNA sequences (HIV-1 TAR-57) using oligonucleotides listed in 4.2.3. The transcription reactions were performed according to the provided protocol at 37 °C for 4 h. Subsequently, 10 units of Calf Intestine Alkaline Phosphatase CIAP (Fermentas, St. Leon-Rot, Germany) were added and the reaction was incubated at 37 °C for another 45 minutes. CIAP dephosphorylates the 5' ends of the newly transcribed RNA molecules and generates compatible ends for radioactive 5' end labeling of the RNA. In order to degrade the DNA template, 10 units of RNase free DNase I (provided in the MEGAscript T7 Kit) were added and the samples were incubated at 37 °C for 20 minutes. The enzymatic reactions were either stopped by storing the samples at -20 °C or by subsequent purification using native polyacrylamide gelelectrophoresis (PAGE). Native PAGE was used to preserve the native folding state of the RNA and to check the homogeneity of the transcription reaction, since in native gels RNA molecules are separated according to their secondary structure. For purification, 6x RNA-sample buffer (final concentration 1x) was added to the samples, which were then loaded immediately onto a gel containing 10 % acrylamide and 1x TBE (for all buffers, see table below). The gel was run in 1x TBE at RT for 45 minutes at 100 V. Afterwards, the RNA bands were visualized using UV shadowing. In case only a single band for each RNA preparation was visible, this band was cut into small pieces and transferred into an eppendorf tube. RNA was extracted from the gel

over night at 37 °C by adding 300 µl crush-and-soak solution. Next, the RNA-buffer solution was separated from the gel pieces by filtration (Amicon Ultrafree-MC centrifugal-filtration devices), the solution was concentrated using centrifugal filter devices (Amicon Ultra) and thereby the crush-and-soak buffer was exchanged with RNase-free water. The concentration of the RNA samples was determined using a NanoDrop spectrophotometer (PEQLAB, Erlangen, Germany) and the samples were stored at -20 °C.

Buffer/Reagent for RNA purification	Composition
10x TBE buffer	0.9 M Tris, 0.9 M boric acid, 20 mM EDTA
gel running buffer	1x TBE in RNase-free H ₂ O
6x RNA-sample buffer	60 mM Tris (pH 7.5), 60 % glycerol
crush-and-soak solution	0.3 M NaCl, 0.5 mM EDTA

4.11.2 Isotopic labeling of RNA oligonucleotides

4.11.2.1 5'-end labeling

To visualize the different RNA oligonucleotides that were used in RNA-binding assays, they were radioactively labeled at their 5' end using γ -³²P-ATP. A 20 µl-labeling reaction contained 10 pmol of the respective RNA oligonucleotide, 15 µCi γ -³²P-ATP, 10 units T4 Polynucleotide Kinase (Fermentas, St. Leon-Rot, Germany), and the supplied buffer A and was carried out for 45 minutes at 37 °C according to the provided protocol. The reaction was stopped by incubation at 70 °C for 10 minutes. Un-incorporated γ -³²P-ATP was removed from the ³²P-labeled oligonucleotides using Sephadex G-25 Quick Spin Columns (Roche Applied Science, Mannheim, Germany). Radiolabeled RNA could be stored at -20 ° for a maximum duration of 2 weeks.

4.11.2.2 3'-end labeling

RNA oligonucleotides of the *EAR1* and *WSC2N* zipcodes were synthesized with a biotin label at their 5' end and thus, had to be radioactively labeled at the 3' end using (5'-³²P)-pCp and T4 RNA ligase. The labeling was carried out in a 20 µl reaction containing 30 pmol of the respective RNA oligonucleotide, 30 pmol (5'-³²P)-pCp, 1 mM ATP, and 10 units T4 RNA ligase (Fermentas, St. Leon-Rot, Germany) in the provided reaction buffer. The reaction was incubated over night at 4 °C, heat inactivated at 70 °C for 10 minutes and finally separated from un-incorporated ³²P-pCp using Sephadex G-25 Quick Spin Columns (Roche Applied Science, Mannheim, Germany). Radiolabeled RNA could be stored at -20 ° for a maximum duration of 2 weeks.

4.11.3 Electrophoretic mobility shift assay (EMSA)

To analyze wild-type She2p binding to different short RNA oligonucleotides (*ASH1* 3-mer to 9-mer motifs, see 4.2.2), EMSAs were performed. In a 20 µl reaction, 2 µM wild-type She2p was incubated with 1 nM of the respective 5' ³²P-labeled RNA oligonucleotide in reaction buffer (20 mM HEPES (pH 7.5), 100 mM NaCl, 1 mM DTT, 5 mM MgCl₂) for 20 minutes at 20 °C. To each sample, 4 µl 6 x RNA-sample buffer (see 4.11.1) was added and the protein:RNA complexes were resolved on a native pre-run gel (7 % acrylamide in 1 x TBE). The gel was run in 1x TBE running buffer at 150 V for 3 h at 20 °C. After drying the gel was exposed to a Storage Phosphor Screen (GE Healthcare, Munich, Germany) for 1 h, scanned with a Storm Scanner (GE Healthcare, Munich, Germany), and analyzed using the ImageQuant software.

4.11.4 RNA filter-binding assay

Besides She2p binding to short RNA oligonucleotides, RNA binding was assessed using a nitrocellulose filter-binding assay (Wong and Lohman, 1993). Serial protein dilutions (see below) and a constant amount of radiolabeled RNA (0.5 nM) were incubated in 20 mM HEPES (pH 7.4), 100 mM NaCl, 2 mM DTT, 2 mM MgCl₂ and 30 µg/ml yeast tRNA at 20 °C for 20 minutes. 80 µl of the reaction mixture was applied on the membrane in a Bio-Dot microfiltration apparatus. After washing and drying, radioactivity retained on the nitrocellulose filter was measured by phosphoimaging using a Storm Scanner (GE Healthcare). Equilibrium-dissociation constants were calculated from a plot of the fraction of bound RNA versus protein concentration using the Langmuir isotherm and the software KaleidaGraph (Synergy Software). For quantification, in each experiment eight serial dilutions were used in the following concentration range: She2p-wt binding to *ASH1* zip-code elements: 0 to 4 µM; She2p-wt binding to HIV-1 TAR, U1snRNA, *EAR1*, and *WSC2N* RNAs: 0 to 12 µM; She2p-wt binding to MS2 RNA: 0 to 6 µM; She2p-mutant binding to bud-localizing RNAs: 0 to 16 µM; She2p-S120Y binding to *ASH1* zipcodes, HIV-1 TAR, or U1snRNA: 0 to 12 µM; She2p(ΔC-term)-binding to *ASH1*-E2A element: 0 to 32 µM; She2p-mutant binding to HIV-1 TAR and U1snRNA: 0 to 32 µM; She2p-wt binding to poly A RNA: 0 to 32 µM.

4.11.5 Surface Plasmon Resonance

Interaction studies of wild type-She2p or the variant She2p-(6-239)-ΔCys with biotinylated RNA oligonucleotides were performed using a Biacore 3000 system (GE Healthcare). Biotinylated RNAs were coupled to a streptavidin chip (GE Healthcare) to signal intensities below 200 response units. In order to obtain sterical flexibility, the 5' biotin tag and the first

nucleobase were separated by an 18-atom spacer (hexaethylene glycol). She2p binding was assessed at 25 °C in running buffer containing 20 mM HEPES (pH 7.5), 150 mM NaCl, and 2 mM DTT. For all experiments, protein concentrations ranging from 1.5 µM to 200 µM were used, with one concentration (25 µM) measured as duplicate. Equilibrium-dissociation constants were obtained from steady-state measurements using the Langmuir isotherm ($R_{eq} = K_A c R_{max} / (1 + c K_A)$, with R_{eq} : steady-state binding level, K_A : association constant, c : analyte concentration, R_{max} : maximum binding level).

4.12 *In vivo* analysis of protein function

4.12.1 Transformation and cultivation of yeast cells

One-step transformation of plasmid DNA was performed as described (Chen et al., 1992). Transformed yeast cells were cultivated on selective media plates for 2 days at 30 °C. Then, transformants were isolated on selective media plates for 1 day at 30 °C. Liquid cultures of yeast strains were grown in YPD or selective medium at 30 °C with 130 rpm. The expression of the protein of interest was confirmed by western blot analysis of yeast whole cell extracts.

4.12.2 Preparation of yeast analytical whole-cell extracts

In order to analyze the protein-expression level of individual transformants, whole-cell extracts were generated. From freshly selected strains, one inoculation loop of cells was resuspended in 1 ml H₂O. Cells were lysed by the addition of 150 µl alkaline solution (1.85 M NaCl, 7.5 % β-mercaptoethanol) and 15 minutes incubation on ice. The cell lysate was precipitated after the addition of 150 µl 50 % trichloroacetic acid for 10 minutes on ice. Following sedimentation (13000 rpm, 15 minutes, 4 °C), the pellet was resuspended in 50 µl HU buffer (8 M urea, 5% SDS, 200 mM Tris pH 8.8, 1 mM EDTA, 1.5% DTT, 0.01 % bromophenol blue) and the proteins were denatured by incubation at 65 °C for 10 minutes. Finally, the samples were analyzed by western blotting for the expression of the desired protein.

4.12.3 Co-immunoprecipitation

Co-immunoprecipitation of myc-tagged She2p and HA-tagged She3p was performed by using monoclonal α-myc antibody 9E11 (Acris Antibodies). 100 µg α-myc antibody has been covalently coupled to 250 µl magnetic Protein-G beads (Invitrogen) according to the manufacturer's instructions using dimethyl pimelimidate (DMP) as a cross linker. 25 µl of

MATERIALS AND METHODS

coupled beads were used for co-immunoprecipitation of a 20 OD culture. Beads were pre-incubated with RNA buffer (for all buffers, see table below) for 3 times 15 minutes to block unspecific RNA binding to the bead surface. 20 OD of an exponentially growing yeast culture were harvested (2000 g, 3 minutes, 4 °C). The pellet was transferred to a 2 ml eppendorf tube and resuspended in 300 µl breaking buffer. 200 µl glass beads (Roth) and 5 µl RiboLock RNase inhibitor (Fermentas) were added and the cells were broken by vortexing 5 times 3 minutes at 4 °C. The lysate was cleared by centrifugation (2000x g, 2 minutes, 4 °C) and the supernatant (250 µl) was transferred to a fresh eppendorf tube. 1/10 of the cleared lysate was taken as input control. The remaining lysate was incubated with 25 µl α -myc coupled beads for 2 h at 4 °C in breaking buffer to allow binding. After binding, the supernatant that contains the unbound fraction was collected and the beads were washed 3 times with wash buffer. Proteins were eluted from the beads by boiling at 95 °C for 10 minutes in 1x Laemmli buffer. 1/50 (v/v) of the input (I) and unbound (FI) fraction, and 1/3 (v/v) of the elution fraction (IP), respectively, were subjected to western-blot analysis. For detection of She2p-3Myc and 3HA-She3p, α -myc 9E10 and α -HA 3F10 antibodies were used, respectively. For quantification, western blots of two independent experiments were analyzed using the LAS-3000 mini system (FUJIFILM). Signal intensities were quantified using the MultiGauge software (FUJIFILM). In all experiments, background signals were determined from $\Delta she2$ lanes and subtracted from respective western-blot signals. Subsequently, signals from the unbound fraction (FI) were subtracted from the input fractions (I) and used to normalize co-immunoprecipitated She3p against the bead-bound fraction of She2p-3Myc.

Buffer for co-immunoprecipitation	Composition
RNA buffer	50 mM HEPES/KOH (pH 7.3), 20 mM potassium acetate, 2 mM EDTA (pH 8.0), 0.1 % (v/v) Triton X-100, 5 % (v/v) glycerol, 0.1 mg/ml yeast tRNA
breaking buffer	50 mM HEPES/KOH (pH 7.3), 20 mM potassium acetate, 2 mM EDTA (pH 8.0), 0.1 % (v/v) Triton X-100, 5 % (v/v) glycerol, 1x Complete EDTA-free protease-inhibitor cocktail (Roche Applied Science)
wash buffer	50 mM HEPES/KOH (pH 7.3), 50 mM potassium acetate, 2 mM magnesium acetate, 0.1 % (v/v) Triton X-100, 5 % (v/v) glycerol
2x Laemmli buffer	125 mM Tris (pH 6.8), 4 % SDS, 200 mM DTT, 20 % glycerol, 0.01 % (w/v) bromophenol blue

4.12.4 Fluorescence microscopy

4.12.4.1 Preparation of yeast spheroplasts

To 15 ml yeast cells of a logarithmically growing culture (OD_{600} 0.8) formaldehyde was added to a final concentration of 3.7 %. The cells were fixed by shaking the culture at 30 °C for another 60 minutes. Cells were harvested (1500x g, 3 minutes) and washed 3 times with spheroplasting premix (1.2 M sorbitol, 0.1 M potassium phosphate (pH 7.4), 0.5 mM $MgCl_2$). Subsequently, the cell pellet was resuspended in 500 μ l spheroplasting solution (1.2 M sorbitol, 0.1 M potassium phosphate (pH 7.4), 0.5 mM $MgCl_2$, 100 μ g/ml Zymolyase 100T, 0.2 % β -mercaptoethanol) and spheroplasting was performed in a 30 °C water bath for 45 minutes. Spheroplasts were briefly pelleted (1000x g, 2 minutes) and washed once in spheroplasting premix. Finally, the cells were resuspended in 150 μ l spheroplasting premix and stored in 50 μ l aliquots at -80 °C.

4.12.4.2 Indirect immunofluorescence

Multi-well slides (Roth) used for indirect immunofluorescence microscopy against wild-type She2p and She2p mutants were coated with 0.02 % poly-L-lysine for 5 minutes and washed 3 times with water. 15 μ l of a spheroplasted cell suspension was applied to each well and incubated for 5 minutes at RT. Then, the cells were incubated with blocking solution (1x PBS, 0.1 % BSA) for 5 minutes. Following, 30 μ l of the primary-antibody solution (α -She2p (323/4), 1:1000 diluted in 1x PBS, 0.1 % BSA) was added to each well and incubated for 2 h at RT in a wet box. After 3 washing steps with washing buffer (1x PBS, 0.1 % BSA, 0.1 % Triton X-100), the secondary-antibody solution was added (Alexa®488 α -rabbit-IgG, 1:250 diluted 1x PBS, 0.1 % BSA) and incubated for 1 h in a darkened wet box. Wells were washed again 3 times, the nuclei were stained with Hoechst Stain Solution (SIGMA), and the cells were mounted in mounting solution (1x PBS, 80 % glycerol). Cells were inspected using an Olympus BX60 fluorescence microscope (Olympus) with a 100x NA 1.3 DIC oil objective and images were acquired by using Openlab 4.01 software (Improvision).

4.13 Methods for structural analysis

4.13.1 NMR data acquisition

1H - ^{15}N HSQC (heteronuclear single quantum coherence) spectra of ^{15}N -labeled She2p were acquired and analyzed by the group of Prof. Michael Sattler (EMBL Heidelberg, Helmholtz

Zentrum München). ^{15}N -labeled She2p was measured at a concentration of 7.1 mg/ml in size-exclusion buffer 5 (see chapter 4.9.4). In the two dimensional ^1H - ^{15}N HSQC experiment, each amino acid (except for proline) of a given protein is described by one peak in the spectrum. For She2p, almost no peaks could be observed. Thus, the planned titration experiments with *ASH1* RNA to map the RNA-binding sites in She2p could not be performed and NMR experiments with She2p were not brought forward.

4.13.2 Crystallization of She2p:RNA complexes

In order to co-crystallize She2p variants with different RNA oligonucleotides (*ASH1*-E3 and variants of *ASH1*-E2B RNA, respectively), 4 mg/ml purified protein in size-exclusion buffer 1 (corresponds to 70 nmol of a She2p dimer) was incubated with a 1.5-molar excess of RNA (dissolved in 20 mM HEPES pH 7.8) for 20 minutes on ice. For initial screening, commercially available screens from Hampton Research (Aliso Viejo, USA), Qiagen (Hilden, Germany), and Jena Bioscience (Jena, Germany) were used. The protein:RNA complex was crystallized at 20 °C by the sitting-drop vapor diffusion technique in 96-well plates with a total reservoir volume of 50 μl by mixing 0.5 μl protein:RNA solution with 0.5 μl of precipitant solution. Optimization of crystallization conditions was performed at 20 °C using the sitting-drop vapor diffusion technique in 24-well plates with a total reservoir volume of 300 μl by mixing 1 μl protein:RNA solution with 1 μl of precipitant solution. In order to test the crystals for diffraction, crystals were incubated in different cryoprotectants (serial transfer from 7.5 % to 15 % ethylene glycol, 5 % to 20 % glycerol, or 15 % to 30 % PEG 400), subsequently mounted in nylon loops, and flash frozen in liquid nitrogen. Crystals were tested at several beamlines at DESY (Hamburg, Germany) and at the ESRF (Grenoble, France), but no data sets were collected.

4.13.3 Small Angle X-ray Scattering (SAXS) experiments and data processing

Synchrotron SAXS data were collected at the X33 beamline (EMBL/DESY, Hamburg, Germany) and at the ID14-3 beamline (ESRF, Grenoble, France). Wild-type She2p and She2p-L130Y were measured in size-exclusion buffer 4 (see chapter 4.9.4) at three concentrations each (1.37 mg/ml, 2.8 mg/ml, 5.92 mg/ml for She2p-wt, and 2.0 mg/ml, 3.75 mg/ml, 6.9 mg/ml for She2p-L130Y). Scattering patterns were measured with exposure times of 2 min (X-33) and ten times for 30 seconds (ID14-3), respectively. Data evaluation and processing was performed using different programs from the ATSAS 2.1 software package (Konarev et al., 2006). PRIMUS was used for primary analysis of the data and to exclude potential sample aggregation during the measurement. Divergence of the normalized

scattering curves at small s -values would be a sign of concentration-dependent protein aggregation or of attractive forces between the molecules. The radius of gyration R_g was calculated using the Guinier approximation with the constraint $s \cdot R_g < 1.3$ (Guinier and Fournet, 1955). The pair-distribution function $P(r)$ and the maximum dimension of the particle D_{\max} were computed using the indirect transform package GNOM. The correct D_{\max} was iteratively determined by evaluating the resulting R_g value and the shape of the $P(r)$ function. The theoretical scattering curves of the She2p-crystal structure (PDB ID: 1XLY) and of modeled tetrameric She2p structures were determined with the program CRY SOL (Svergun et al., 1995). SAXS data were collected for wild-type She2p and She2p-L130Y in various buffer conditions. Except for protein samples that were measured in size-exclusion buffer 4, we observed interparticle interference or protein aggregation at higher concentrations. Thus, only data were processed that were collected for both proteins in size-exclusion buffer 4. As reference proteins for determination of molecular masses (Mylonas and Svergun, 2007), bovine serum albumine (5.7 mg/ml) and lysozyme (8.5 mg/ml) were measured in buffer containing 50 mM HEPES (pH 7.5) and 2 mM DTT. *Ab initio* models of wild-type She2p were calculated with the programs GASBORp and GASBORi (Svergun et al., 2001). GASBORp uses the pair-distribution function for modeling, whereas GASBORi uses the non-transformed scattering intensity. For modeling, different particle symmetries were assumed, but only calculations without any given symmetry (P1) gave reasonable results. Since at the highest protein concentration, the quality of scattering data was reduced, several bead models were calculated for different protein concentrations. For each final bead model, 10 identically calculated models were averaged and aligned to the She2p structure using the programs DAMAVER and SUPCOMB (Volkov and Svergun, 2003).

5 References

- Ainger, K., Avossa, D., Diana, A.S., Barry, C., Barbarese, E., and Carson, J.H. (1997). Transport and localization elements in myelin basic protein mRNA. *J Cell Biol* *138*, 1077-1087.
- Allen, L., Kloc, M., and Etkin, L.D. (2003). Identification and characterization of the Xlsirt cis-acting RNA localization element. *Differentiation* *71*, 311-321.
- Angenstein, F., Evans, A.M., Ling, S.C., Settlage, R.E., Ficarro, S., Carrero-Martinez, F.A., Shabanowitz, J., Hunt, D.F., and Greenough, W.T. (2005). Proteomic characterization of messenger ribonucleoprotein complexes bound to nontranslated or translated poly(A) mRNAs in the rat cerebral cortex. *J Biol Chem* *280*, 6496-6503.
- Atlas, R., Behar, L., Elliott, E., and Ginzburg, I. (2004). The insulin-like growth factor mRNA binding-protein IMP-1 and the Ras-regulatory protein G3BP associate with tau mRNA and HuD protein in differentiated P19 neuronal cells. *J Neurochem* *89*, 613-626.
- Auweter, S.D., Oberstrass, F.C., and Allain, F.H. (2006). Sequence-specific binding of single-stranded RNA: is there a code for recognition? *Nucleic Acids Res* *34*, 4943-4959.
- Bashirullah, A., Halsell, S.R., Cooperstock, R.L., Kloc, M., Karauskakis, A., Fisher, W.W., Fu, W., Hamilton, J.K., Etkin, L.D., and Lipshitz, H.D. (1999). Joint action of two RNA degradation pathways controls the timing of maternal transcript elimination at the midblastula transition in *Drosophila melanogaster*. *EMBO J* *18*, 2610-2620.
- Beach, D.L., Salmon, E.D., and Bloom, K. (1999). Localization and anchoring of mRNA in budding yeast. *Curr Biol* *9*, 569-578.
- Berleth, T., Burri, M., Thoma, G., Bopp, D., Richstein, S., Frigerio, G., Noll, M., and Nusslein-Volhard, C. (1988). The role of localization of bicoid RNA in organizing the anterior pattern of the *Drosophila* embryo. *EMBO J* *7*, 1749-1756.
- Bertrand, E., Chartrand, P., Schaefer, M., Shenoy, S.M., Singer, R.H., and Long, R.M. (1998). Localization of ASH1 mRNA particles in living yeast. *Mol Cell* *2*, 437-445.
- Besse, F., and Ephrussi, A. (2008). Translational control of localized mRNAs: restricting protein synthesis in space and time. *Nat Rev Mol Cell Biol* *9*, 971-980.
- Besse, F., Lopez de Quinto, S., Marchand, V., Trucco, A., and Ephrussi, A. (2009). *Drosophila* PTB promotes formation of high-order RNP particles and represses oskar translation. *Genes Dev* *23*, 195-207.
- Beuth, B., Pennell, S., Arnvig, K.B., Martin, S.R., and Taylor, I.A. (2005). Structure of a *Mycobacterium tuberculosis* NusA-RNA complex. *EMBO J* *24*, 3576-3587.

REFERENCES

- Blobel, G., and Dobberstein, B. (1975). Transfer of proteins across membranes. I. Presence of proteolytically processed and unprocessed nascent immunoglobulin light chains on membrane-bound ribosomes of murine myeloma. *J Cell Biol* *67*, 835-851.
- Bobola, N., Jansen, R.P., Shin, T.H., and Nasmyth, K. (1996). Asymmetric accumulation of Ash1p in postanaphase nuclei depends on a myosin and restricts yeast mating-type switching to mother cells. *Cell* *84*, 699-709.
- Böhl, F., Kruse, C., Frank, A., Ferring, D., and Jansen, R.P. (2000). She2p, a novel RNA-binding protein tethers ASH1 mRNA to the Myo4p myosin motor via She3p. *EMBO J* *19*, 5514-5524.
- Brenner, H.R., Witzemann, V., and Sakmann, B. (1990). Imprinting of acetylcholine receptor messenger RNA accumulation in mammalian neuromuscular synapses. *Nature* *344*, 544-547.
- Bullock, S.L. (2007). Translocation of mRNAs by molecular motors: think complex? *Semin Cell Dev Biol* *18*, 194-201.
- Bullock, S.L., and Ish-Horowicz, D. (2001). Conserved signals and machinery for RNA transport in *Drosophila* oogenesis and embryogenesis. *Nature* *414*, 611-616.
- Burd, C.G., and Dreyfuss, G. (1994). Conserved structures and diversity of functions of RNA-binding proteins. *Science* *265*, 615-621.
- Chang, P., Torres, J., Lewis, R.A., Mowry, K.L., Houliston, E., and King, M.L. (2004). Localization of RNAs to the mitochondrial cloud in *Xenopus* oocytes through entrapment and association with endoplasmic reticulum. *Mol Biol Cell* *15*, 4669-4681.
- Chartrand, P., Meng, X.H., Hüttelmaier, S., Donato, D., and Singer, R.H. (2002). Asymmetric sorting of ash1p in yeast results from inhibition of translation by localization elements in the mRNA. *Mol Cell* *10*, 1319-1330.
- Chartrand, P., Meng, X.H., Singer, R.H., and Long, R.M. (1999). Structural elements required for the localization of ASH1 mRNA and of a green fluorescent protein reporter particle in vivo. *Curr Biol* *9*, 333-336.
- Chekulaeva, M., Hentze, M.W., and Ephrussi, A. (2006). Bruno acts as a dual repressor of oskar translation, promoting mRNA oligomerization and formation of silencing particles. *Cell* *124*, 521-533.
- Chen, D.C., Yang, B.C., and Kuo, T.T. (1992). One-step transformation of yeast in stationary phase. *Curr Genet* *21*, 83-84.
- Condeelis, J., and Singer, R.H. (2005). How and why does beta-actin mRNA target? *Biol Cell* *97*, 97-110.
- Cosma, M.P. (2004). Daughter-specific repression of *Saccharomyces cerevisiae* HO: Ash1 is the commander. *EMBO Rep* *5*, 953-957.

REFERENCES

- Dahm, R., and Kiebler, M. (2005). Cell biology: silenced RNA on the move. *Nature* *438*, 432-435.
- Dahm, R., Kiebler, M., and Macchi, P. (2007). RNA localisation in the nervous system. *Semin Cell Dev Biol* *18*, 216-223.
- Darnell, J.C., Mostovetsky, O., and Darnell, R.B. (2005). FMRP RNA targets: identification and validation. *Genes Brain Behav* *4*, 341-349.
- DeLano, W.L. (2004). Use of PYMOL as a communications tool for molecular science. *Abstracts of Papers of the American Chemical Society* *228*, U313-U314.
- Demeler, B. (2005). UltraScan A Comprehensive Data Analysis Software Package for Analytical Ultracentrifugation Experiments. in *Modern Analytical Ultracentrifugation: Techniques and Methods*, 210-229.
- Deng, Y., Singer, R.H., and Gu, W. (2008). Translation of ASH1 mRNA is repressed by Puf6p-Fun12p/eIF5B interaction and released by CK2 phosphorylation. *Genes Dev* *22*, 1037-1050.
- Deshler, J.O., Highett, M.I., Abramson, T., and Schnapp, B.J. (1998). A highly conserved RNA-binding protein for cytoplasmic mRNA localization in vertebrates. *Curr Biol* *8*, 489-496.
- Deshler, J.O., Highett, M.I., and Schnapp, B.J. (1997). Localization of *Xenopus* Vg1 mRNA by Vera protein and the endoplasmic reticulum. *Science* *276*, 1128-1131.
- Ding, D., Parkhurst, S.M., Halsell, S.R., and Lipshitz, H.D. (1993). Dynamic Hsp83 RNA localization during *Drosophila* oogenesis and embryogenesis. *Mol Cell Biol* *13*, 3773-3781.
- Dreyfuss, G., Kim, V.N., and Kataoka, N. (2002). Messenger-RNA-binding proteins and the messages they carry. *Nat Rev Mol Cell Biol* *3*, 195-205.
- Du, T.G., Jellbauer, S., Müller, M., Schmid, M., Niessing, D., and Jansen, R.P. (2008). Nuclear transit of the RNA-binding protein She2 is required for translational control of localized ASH1 mRNA. *EMBO Rep* *9*, 781-787.
- Duncan, J.E., and Warrior, R. (2002). The cytoplasmic dynein and kinesin motors have interdependent roles in patterning the *Drosophila* oocyte. *Curr Biol* *12*, 1982-1991.
- Ephrussi, A., Dickinson, L.K., and Lehmann, R. (1991). Oskar organizes the germ plasm and directs localization of the posterior determinant nanos. *Cell* *66*, 37-50.
- Ephrussi, A., and St Johnston, D. (2004). Seeing is believing: the bicoid morphogen gradient matures. *Cell* *116*, 143-152.
- Estrada, P., Kim, J., Coleman, J., Walker, L., Dunn, B., Takizawa, P., Novick, P., and Ferro-Novick, S. (2003). Myo4p and She3p are required for cortical ER inheritance in *Saccharomyces cerevisiae*. *J Cell Biol* *163*, 1255-1266.

REFERENCES

- Evangelista, M., Zigmond, S., and Boone, C. (2003). Formins: signaling effectors for assembly and polarization of actin filaments. *J Cell Sci* *116*, 2603-2611.
- Evdokimova, V., Ruzanov, P., Anglesio, M.S., Sorokin, A.V., Ovchinnikov, L.P., Buckley, J., Triche, T.J., Sonenberg, N., and Sorensen, P.H. (2006). Akt-mediated YB-1 phosphorylation activates translation of silent mRNA species. *Mol Cell Biol* *26*, 277-292.
- Farina, K.L., Hüttelmaier, S., Musunuru, K., Darnell, R., and Singer, R.H. (2003). Two ZBP1 KH domains facilitate beta-actin mRNA localization, granule formation, and cytoskeletal attachment. *J Cell Biol* *160*, 77-87.
- Ferrandon, D., Elphick, L., Nusslein-Volhard, C., and St Johnston, D. (1994). Stauf protein associates with the 3' UTR of bicoid mRNA to form particles that move in a microtubule-dependent manner. *Cell* *79*, 1221-1232.
- Forrest, K.M., and Gavis, E.R. (2003). Live imaging of endogenous RNA reveals a diffusion and entrapment mechanism for nanos mRNA localization in *Drosophila*. *Curr Biol* *13*, 1159-1168.
- Furic, L., Maher-Laporte, M., and DesGroseillers, L. (2008). A genome-wide approach identifies distinct but overlapping subsets of cellular mRNAs associated with Stauf1- and Stauf2-containing ribonucleoprotein complexes. *RNA* *14*, 324-335.
- Gasteiger, E., Gattiker, A., Hoogland, C., Ivanyi, I., Appel, R.D., and Bairoch, A. (2003). ExPASy: The proteomics server for in-depth protein knowledge and analysis. *Nucleic Acids Res* *31*, 3784-3788.
- Gavis, E.R., Curtis, D., and Lehmann, R. (1996). Identification of cis-acting sequences that control nanos RNA localization. *Dev Biol* *176*, 36-50.
- Gerber, A.P., Herschlag, D., and Brown, P.O. (2004). Extensive association of functionally and cytologically related mRNAs with Puf family RNA-binding proteins in yeast. *PLoS Biol* *2*, E79.
- Gietz, R.D., and Sugino, A. (1988). New yeast-*Escherichia coli* shuttle vectors constructed with in vitro mutagenized yeast genes lacking six-base pair restriction sites. *Gene* *74*, 527-534.
- Giorgi, C., and Moore, M.J. (2007). The nuclear nurture and cytoplasmic nature of localized mRNPs. *Semin Cell Dev Biol* *18*, 186-193.
- Gonsalvez, G.B., Lehmann, K.A., Ho, D.K., Stanitsa, E.S., Williamson, J.R., and Long, R.M. (2003). RNA-protein interactions promote asymmetric sorting of the ASH1 mRNA ribonucleoprotein complex. *RNA* *9*, 1383-1399.
- Gonsalvez, G.B., Little, J.L., and Long, R.M. (2004). ASH1 mRNA anchoring requires reorganization of the Myo4p-She3p-She2p transport complex. *J Biol Chem* *279*, 46286-46294.

REFERENCES

- Gonzalez-Reyes, A., Elliott, H., and St Johnston, D. (1995). Polarization of both major body axes in *Drosophila* by gurken-torpedo signalling. *Nature* *375*, 654-658.
- Gonzalez, I., Buonomo, S.B., Nasmyth, K., and von Ahsen, U. (1999). ASH1 mRNA localization in yeast involves multiple secondary structural elements and Ash1 protein translation. *Curr Biol* *9*, 337-340.
- Gu, W., Deng, Y., Zenklusen, D., and Singer, R.H. (2004). A new yeast PUF family protein, Puf6p, represses ASH1 mRNA translation and is required for its localization. *Genes Dev* *18*, 1452-1465.
- Gu, W., Pan, F., Zhang, H., Bassell, G.J., and Singer, R.H. (2002). A predominantly nuclear protein affecting cytoplasmic localization of beta-actin mRNA in fibroblasts and neurons. *J Cell Biol* *156*, 41-51.
- Guinier, A., and Fournet, F. (1955). Small Angle Scattering of X-rays. Wiley Interscience, New York.
- Hachet, O., and Ephrussi, A. (2004). Splicing of oskar RNA in the nucleus is coupled to its cytoplasmic localization. *Nature* *428*, 959-963.
- Hamilton, R.S., and Davis, I. (2007). RNA localization signals: deciphering the message with bioinformatics. *Semin Cell Dev Biol* *18*, 178-185.
- Harnpicharnchai, P., Jakovljevic, J., Horsey, E., Miles, T., Roman, J., Rout, M., Meagher, D., Imai, B., Guo, Y., Brame, C.J., *et al.* (2001). Composition and functional characterization of yeast 66S ribosome assembly intermediates. *Mol Cell* *8*, 505-515.
- Havin, L., Git, A., Elisha, Z., Oberman, F., Yaniv, K., Schwartz, S.P., Standart, N., and Yisraeli, J.K. (1998). RNA-binding protein conserved in both microtubule- and microfilament-based RNA localization. *Genes Dev* *12*, 1593-1598.
- Heinrich, B., and Deshler, J.O. (2009). RNA localization to the Balbiani body in *Xenopus* oocytes is regulated by the energy state of the cell and is facilitated by kinesin II. *RNA*.
- Heuck, A., Du, T.G., Jellbauer, S., Richter, K., Kruse, C., Jaklin, S., Müller, M., Buchner, J., Jansen, R.P., and Niessing, D. (2007). Monomeric myosin V uses two binding regions for the assembly of stable translocation complexes. *Proc Natl Acad Sci U S A* *104*, 19778-19783.
- Ho, S.N., Hunt, H.D., Horton, R.M., Pullen, J.K., and Pease, L.R. (1989). Site-directed mutagenesis by overlap extension using the polymerase chain reaction. *Gene* *77*, 51-59.
- Hoek, K.S., Kidd, G.J., Carson, J.H., and Smith, R. (1998). hnRNP A2 selectively binds the cytoplasmic transport sequence of myelin basic protein mRNA. *Biochemistry* *37*, 7021-7029.

REFERENCES

- Hogan, D.J., Riordan, D.P., Gerber, A.P., Herschlag, D., and Brown, P.O. (2008). Diverse RNA-binding proteins interact with functionally related sets of RNAs, suggesting an extensive regulatory system. *PLoS Biol* 6, e255.
- Hook, B.A., Goldstrohm, A.C., Seay, D.J., and Wickens, M. (2007). Two yeast PUF proteins negatively regulate a single mRNA. *J Biol Chem* 282, 15430-15438.
- Huang, F., Chotiner, J.K., and Steward, O. (2007). Actin polymerization and ERK phosphorylation are required for Arc/Arg3.1 mRNA targeting to activated synaptic sites on dendrites. *J Neurosci* 27, 9054-9067.
- Hurt, E., Strasser, K., Segref, A., Bailer, S., Schlaich, N., Presutti, C., Tollervey, D., and Jansen, R. (2000). Mex67p mediates nuclear export of a variety of RNA polymerase II transcripts. *J Biol Chem* 275, 8361-8368.
- Hüttelmaier, S., Zenklusen, D., Lederer, M., Dichtenberg, J., Lorenz, M., Meng, X., Bassell, G.J., Condeelis, J., and Singer, R.H. (2005). Spatial regulation of beta-actin translation by Src-dependent phosphorylation of ZBP1. *Nature* 438, 512-515.
- Irie, K., Tadauchi, T., Takizawa, P.A., Vale, R.D., Matsumoto, K., and Herskowitz, I. (2002). The Khd1 protein, which has three KH RNA-binding motifs, is required for proper localization of ASH1 mRNA in yeast. *EMBO J* 21, 1158-1167.
- Jambhekar, A., and DeRisi, J.L. (2007). Cis-acting determinants of asymmetric, cytoplasmic RNA transport. *RNA* 13, 625-642.
- Jambhekar, A., McDermott, K., Sorber, K., Shepard, K.A., Vale, R.D., Takizawa, P.A., and DeRisi, J.L. (2005). Unbiased selection of localization elements reveals cis-acting determinants of mRNA bud localization in *Saccharomyces cerevisiae*. *Proc Natl Acad Sci U S A* 102, 18005-18010.
- Jansen, R.P., Dowzer, C., Michaelis, C., Galova, M., and Nasmyth, K. (1996). Mother cell-specific HO expression in budding yeast depends on the unconventional myosin myo4p and other cytoplasmic proteins. *Cell* 84, 687-697.
- Januschke, J., Gervais, L., Dass, S., Kaltschmidt, J.A., Lopez-Schier, H., St Johnston, D., Brand, A.H., Roth, S., and Guichet, A. (2002). Polar transport in the *Drosophila* oocyte requires Dynein and Kinesin I cooperation. *Curr Biol* 12, 1971-1981.
- Job, C., and Eberwine, J. (2001). Localization and translation of mRNA in dendrites and axons. *Nat Rev Neurosci* 2, 889-898.
- Kanai, Y., Dohmae, N., and Hirokawa, N. (2004). Kinesin transports RNA: isolation and characterization of an RNA-transporting granule. *Neuron* 43, 513-525.
- Kiebler, M.A., and Bassell, G.J. (2006). Neuronal RNA granules: movers and makers. *Neuron* 51, 685-690.

REFERENCES

- Kim-Ha, J., Kerr, K., and Macdonald, P.M. (1995). Translational regulation of oskar mRNA by bruno, an ovarian RNA-binding protein, is essential. *Cell* *81*, 403-412.
- Kim-Ha, J., Smith, J.L., and Macdonald, P.M. (1991). oskar mRNA is localized to the posterior pole of the *Drosophila* oocyte. *Cell* *66*, 23-35.
- Kislauskis, E.H., Zhu, X., and Singer, R.H. (1994). Sequences responsible for intracellular localization of beta-actin messenger RNA also affect cell phenotype. *J Cell Biol* *127*, 441-451.
- Kloc, M., Bilinski, S., Pui-Yee Chan, A., and Etkin, L.D. (2000). The targeting of Xcat2 mRNA to the germinal granules depends on a cis-acting germinal granule localization element within the 3' UTR. *Dev Biol* *217*, 221-229.
- Kloc, M., and Etkin, L.D. (1995). Two distinct pathways for the localization of RNAs at the vegetal cortex in *Xenopus* oocytes. *Development* *121*, 287-297.
- Kloc, M., and Etkin, L.D. (1998). Apparent continuity between the messenger transport organizer and late RNA localization pathways during oogenesis in *Xenopus*. *Mech Dev* *73*, 95-106.
- Kloc, M., and Etkin, L.D. (2005). RNA localization mechanisms in oocytes. *J Cell Sci* *118*, 269-282.
- Kloc, M., Larabell, C., and Etkin, L.D. (1996). Elaboration of the messenger transport organizer pathway for localization of RNA to the vegetal cortex of *Xenopus* oocytes. *Dev Biol* *180*, 119-130.
- Kloc, M., Spohr, G., and Etkin, L.D. (1993). Translocation of repetitive RNA sequences with the germ plasm in *Xenopus* oocytes. *Science* *262*, 1712-1714.
- Kloc, M., Zearfoss, N.R., and Etkin, L.D. (2002). Mechanisms of subcellular mRNA localization. *Cell* *108*, 533-544.
- Koh, Y.Y., Opperman, L., Stumpf, C., Mandan, A., Keles, S., and Wickens, M. (2009). A single *C. elegans* PUF protein binds RNA in multiple modes. *RNA*.
- Komili, S., Farny, N.G., Roth, F.P., and Silver, P.A. (2007). Functional specificity among ribosomal proteins regulates gene expression. *Cell* *131*, 557-571.
- Konarev, P.V., Petoukhov, M.V., Volkov, V.V., and Svergun, D.I. (2006). ATSAS 2.1, a program package for small-angle scattering data analysis. *Journal of Applied Crystallography* *39*, 277-286.
- Kruse, C., Jaedicke, A., Beaudouin, J., Böhl, F., Ferring, D., Guttler, T., Ellenberg, J., and Jansen, R.P. (2002). Ribonucleoprotein-dependent localization of the yeast class V myosin Myo4p. *J Cell Biol* *159*, 971-982.
- Laemmli, U.K. (1970). Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature* *227*, 680-685.

REFERENCES

- Lange, S., Katayama, Y., Schmid, M., Burkacky, O., Brauchle, C., Lamb, D.C., and Jansen, R.P. (2008). Simultaneous transport of different localized mRNA species revealed by live-cell imaging. *Traffic* *9*, 1256-1267.
- Lawrence, J.B., and Singer, R.H. (1986). Intracellular localization of messenger RNAs for cytoskeletal proteins. *Cell* *45*, 407-415.
- Lecuyer, E., Yoshida, H., Parthasarathy, N., Alm, C., Babak, T., Cerovina, T., Hughes, T.R., Tomancak, P., and Krause, H.M. (2007). Global analysis of mRNA localization reveals a prominent role in organizing cellular architecture and function. *Cell* *131*, 174-187.
- Lewis, R.A., Kress, T.L., Cote, C.A., Gautreau, D., Rokop, M.E., and Mowry, K.L. (2004). Conserved and clustered RNA recognition sequences are a critical feature of signals directing RNA localization in *Xenopus* oocytes. *Mech Dev* *121*, 101-109.
- Long, R.M., Gu, W., Lorimer, E., Singer, R.H., and Chartrand, P. (2000). She2p is a novel RNA-binding protein that recruits the Myo4p-She3p complex to ASH1 mRNA. *EMBO J* *19*, 6592-6601.
- Long, R.M., Gu, W., Meng, X., Gonsalvez, G., Singer, R.H., and Chartrand, P. (2001). An exclusively nuclear RNA-binding protein affects asymmetric localization of ASH1 mRNA and Ash1p in yeast. *J Cell Biol* *153*, 307-318.
- Long, R.M., Singer, R.H., Meng, X., Gonzalez, I., Nasmyth, K., and Jansen, R.P. (1997). Mating type switching in yeast controlled by asymmetric localization of ASH1 mRNA. *Science* *277*, 383-387.
- Lunde, B.M., Moore, C., and Varani, G. (2007). RNA-binding proteins: modular design for efficient function. *Nat Rev Mol Cell Biol* *8*, 479-490.
- Macchi, P., Brownawell, A.M., Grunewald, B., DesGroseillers, L., Macara, I.G., and Kiebler, M.A. (2004). The brain-specific double-stranded RNA-binding protein Staufen2: nucleolar accumulation and isoform-specific exportin-5-dependent export. *J Biol Chem* *279*, 31440-31444.
- Martin, K.C., and Ephrussi, A. (2009). mRNA localization: gene expression in the spatial dimension. *Cell* *136*, 719-730.
- Messitt, T.J., Gagnon, J.A., Kreiling, J.A., Pratt, C.A., Yoon, Y.J., and Mowry, K.L. (2008). Multiple kinesin motors coordinate cytoplasmic RNA transport on a subpopulation of microtubules in *Xenopus* oocytes. *Dev Cell* *15*, 426-436.
- Micklem, D.R., Adams, J., Grunert, S., and St Johnston, D. (2000). Distinct roles of two conserved Staufen domains in oskar mRNA localization and translation. *EMBO J* *19*, 1366-1377.
- Miller, M.T., Higgin, J.J., and Hall, T.M. (2008). Basis of altered RNA-binding specificity by PUF proteins revealed by crystal structures of yeast Puf4p. *Nat Struct Mol Biol* *15*, 397-402.

REFERENCES

- Münchow, S., Sauter, C., and Jansen, R.P. (1999). Association of the class V myosin Myo4p with a localised messenger RNA in budding yeast depends on She proteins. *J Cell Sci* 112 (Pt 10), 1511-1518.
- Munro, T.P., Magee, R.J., Kidd, G.J., Carson, J.H., Barbarese, E., Smith, L.M., and Smith, R. (1999). Mutational analysis of a heterogeneous nuclear ribonucleoprotein A2 response element for RNA trafficking. *J Biol Chem* 274, 34389-34395.
- Mylonas, E., and Svergun, D.I. (2007). Accuracy of molecular mass determination of proteins in solution by small-angle X-ray scattering. *J Appl Cryst* 40, 245-249.
- Nakamura, A., Sato, K., and Hanyu-Nakamura, K. (2004). *Drosophila* cup is an eIF4E binding protein that associates with Bruno and regulates oskar mRNA translation in oogenesis. *Dev Cell* 6, 69-78.
- Neuman-Silberberg, F.S., and Schupbach, T. (1993). The *Drosophila* dorsoventral patterning gene *gurken* produces a dorsally localized RNA and encodes a TGF alpha-like protein. *Cell* 75, 165-174.
- Niessing, D., Hüttelmaier, S., Zenklusen, D., Singer, R.H., and Burley, S.K. (2004). She2p is a novel RNA binding protein with a basic helical hairpin motif. *Cell* 119, 491-502.
- Oeffinger, M., Wei, K.E., Rogers, R., DeGrasse, J.A., Chait, B.T., Aitchison, J.D., and Rout, M.P. (2007). Comprehensive analysis of diverse ribonucleoprotein complexes. *Nat Methods* 4, 951-956.
- Ohashi, S., Koike, K., Omori, A., Ichinose, S., Ohara, S., Kobayashi, S., Sato, T.A., and Anzai, K. (2002). Identification of mRNA/protein (mRNP) complexes containing Puralpha, mStaufen, fragile X protein, and myosin Va and their association with rough endoplasmic reticulum equipped with a kinesin motor. *J Biol Chem* 277, 37804-37810.
- Olivier, C., Poirier, G., Gendron, P., Boisgontier, A., Major, F., and Chartrand, P. (2005). Identification of a conserved RNA motif essential for She2p recognition and mRNA localization to the yeast bud. *Mol Cell Biol* 25, 4752-4766.
- Pan, F., Hüttelmaier, S., Singer, R.H., and Gu, W. (2007). ZBP2 facilitates binding of ZBP1 to beta-actin mRNA during transcription. *Mol Cell Biol* 27, 8340-8351.
- Paquin, N., and Chartrand, P. (2008). Local regulation of mRNA translation: new insights from the bud. *Trends Cell Biol* 18, 105-111.
- Paquin, N., Menade, M., Poirier, G., Donato, D., Drouet, E., and Chartrand, P. (2007). Local activation of yeast ASH1 mRNA translation through phosphorylation of Khd1p by the casein kinase Yck1p. *Mol Cell* 26, 795-809.
- Parisien, M., and Major, F. (2008). The MC-Fold and MC-Sym pipeline infers RNA structure from sequence data. *Nature* 452, 51-55.

REFERENCES

- Perez-Canadillas, J.M. (2006). Grabbing the message: structural basis of mRNA 3' UTR recognition by Hrp1. *EMBO J* *25*, 3167-3178.
- Ponchon, L., and Dardel, F. (2007). Recombinant RNA technology: the tRNA scaffold. *Nat Methods* *4*, 571-576.
- Putnam, C.D., Hammel, M., Hura, G.L., and Tainer, J.A. (2007). X-ray solution scattering (SAXS) combined with crystallography and computation: defining accurate macromolecular structures, conformations and assemblies in solution. *Q Rev Biophys* *40*, 191-285.
- Ramos, A., Bayer, P., and Varani, G. (1999). Determination of the structure of the RNA complex of a double-stranded RNA-binding domain from *Drosophila* Stauf protein. *Biopolymers* *52*, 181-196.
- Ramos, A., Grunert, S., Adams, J., Micklem, D.R., Proctor, M.R., Freund, S., Bycroft, M., St Johnston, D., and Varani, G. (2000). RNA recognition by a Stauf double-stranded RNA-binding domain. *EMBO J* *19*, 997-1009.
- Ross, A.F., Oleynikov, Y., Kislauskis, E.H., Taneja, K.L., and Singer, R.H. (1997). Characterization of a beta-actin mRNA zipcode-binding protein. *Mol Cell Biol* *17*, 2158-2165.
- Sambrook, J., and Russell, D.W. (2001). *Molecular cloning - A Laboratory Manual*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York.
- Schmid, M., Jaedicke, A., Du, T.G., and Jansen, R.P. (2006). Coordination of endoplasmic reticulum and mRNA localization to the yeast bud. *Curr Biol* *16*, 1538-1543.
- Shen, Z., Paquin, N., Forget, A., and Chartrand, P. (2009). Nuclear Shuttling of She2p Couples ASH1 mRNA Localization to its Translational Repression by Recruiting Loc1p and Puf6p. *Mol Biol Cell*.
- Shepard, K.A., Gerber, A.P., Jambhekar, A., Takizawa, P.A., Brown, P.O., Herschlag, D., DeRisi, J.L., and Vale, R.D. (2003). Widespread cytoplasmic mRNA transport in yeast: identification of 22 bud-localized transcripts using DNA microarray analysis. *Proc Natl Acad Sci U S A* *100*, 11429-11434.
- Sil, A., and Herskowitz, I. (1996). Identification of asymmetrically localized determinant, Ash1p, required for lineage-specific transcription of the yeast HO gene. *Cell* *84*, 711-722.
- Simon, A.M., Hoppe, P., and Burden, S.J. (1992). Spatial restriction of AChR gene expression to subsynaptic nuclei. *Development* *114*, 545-553.
- Snee, M., Benz, D., Jen, J., and Macdonald, P.M. (2008). Two distinct domains of Bruno bind specifically to the oskar mRNA. *RNA Biol* *5*, 1-9.
- St Johnston, D. (1995). The intracellular localization of messenger RNAs. *Cell* *81*, 161-170.

REFERENCES

- St Johnston, D. (2005). Moving messages: the intracellular localization of mRNAs. *Nat Rev Mol Cell Biol* *6*, 363-375.
- St Johnston, D., Beuchle, D., and Nusslein-Volhard, C. (1991). *Staufen*, a gene required to localize maternal RNAs in the *Drosophila* egg. *Cell* *66*, 51-63.
- St Johnston, D., Brown, N.H., Gall, J.G., and Jantsch, M. (1992). A conserved double-stranded RNA-binding domain. *Proc Natl Acad Sci U S A* *89*, 10979-10983.
- Svergun, D.I., Barberato, C., and Koch, M.H. (1995). CRY SOL-a Program to Evaluate X-ray Solution Scattering of Biological Macromolecules from Atomic Coordinates *J Appl Cryst* *28*, 768-773.
- Svergun, D.I., Petoukhov, M.V., and Koch, M.H. (2001). Determination of domain structure of proteins from X-ray solution scattering. *Biophys J* *80*, 2946-2953.
- Svoboda, P., and Di Cara, A. (2006). Hairpin RNA: a secondary structure of primary importance. *Cell Mol Life Sci* *63*, 901-908.
- Takizawa, P.A., Sil, A., Swedlow, J.R., Herskowitz, I., and Vale, R.D. (1997). Actin-dependent localization of an RNA encoding a cell-fate determinant in yeast. *Nature* *389*, 90-93.
- Takizawa, P.A., and Vale, R.D. (2000). The myosin motor, Myo4p, binds Ash1 mRNA via the adapter protein, She3p. *Proc Natl Acad Sci U S A* *97*, 5273-5278.
- Tenenbaum, S.A., Carson, C.C., Lager, P.J., and Keene, J.D. (2000). Identifying mRNA subsets in messenger ribonucleoprotein complexes by using cDNA arrays. *Proc Natl Acad Sci U S A* *97*, 14085-14090.
- Tiruchinapalli, D.M., Oleynikov, Y., Kelic, S., Shenoy, S.M., Hartley, A., Stanton, P.K., Singer, R.H., and Bassell, G.J. (2003). Activity-dependent trafficking and dynamic localization of zipcode binding protein 1 and beta-actin mRNA in dendrites and spines of hippocampal neurons. *J Neurosci* *23*, 3251-3261.
- Ule, J., Jensen, K.B., Ruggiu, M., Mele, A., Ule, A., and Darnell, R.B. (2003). CLIP identifies Nova-regulated RNA networks in the brain. *Science* *302*, 1212-1215.
- Urbinati, C.R., Gonsalvez, G.B., Aris, J.P., and Long, R.M. (2006). Loc1p is required for efficient assembly and nuclear export of the 60S ribosomal subunit. *Mol Genet Genomics* *276*, 369-377.
- Vale, R.D. (2003). The molecular motor toolbox for intracellular transport. *Cell* *112*, 467-480.
- Volkov, V.V., and Svergun, D.I. (2003). Uniqueness of ab initio shape determination in small-angle scattering. *Journal of Applied Crystallography* *36*, 860-864.
- Wang, X., McLachlan, J., Zamore, P.D., and Hall, T.M. (2002). Modular recognition of RNA by a human pumilio-homology domain. *Cell* *110*, 501-512.

REFERENCES

- Webster, P.J., Liang, L., Berg, C.A., Lasko, P., and Macdonald, P.M. (1997). Translational repressor bruno plays multiple roles in development and is widely conserved. *Genes Dev* *11*, 2510-2521.
- Weil, T.T., Parton, R., Davis, I., and Gavis, E.R. (2008). Changes in bicoid mRNA anchoring highlight conserved mechanisms during the oocyte-to-embryo transition. *Curr Biol* *18*, 1055-1061.
- Wesche, S., Arnold, M., and Jansen, R.P. (2003). The UCS domain protein She4p binds to myosin motor domains and is essential for class I and class V myosin function. *Curr Biol* *13*, 715-724.
- Wickens, M., Bernstein, D.S., Kimble, J., and Parker, R. (2002). A PUF family portrait: 3' UTR regulation as a way of life. *Trends Genet* *18*, 150-157.
- Wilhelm, J.E., Hilton, M., Amos, Q., and Henzel, W.J. (2003). Cup is an eIF4E binding protein required for both the translational repression of oskar and the recruitment of Barentsz. *J Cell Biol* *163*, 1197-1204.
- Wilhelm, J.E., and Vale, R.D. (1993). RNA on the move: the mRNA localization pathway. *J Cell Biol* *123*, 269-274.
- Wong, I., and Lohman, T.M. (1993). A double-filter method for nitrocellulose-filter binding: application to protein-nucleic acid interactions. *Proc Natl Acad Sci U S A* *90*, 5428-5432.
- Worbs, M., Bourenkov, G.P., Bartunik, H.D., Huber, R., and Wahl, M.C. (2001). An extended RNA binding surface through arrayed S1 and KH domains in transcription factor NusA. *Mol Cell* *7*, 1177-1189.
- Yisraeli, J.K., Sokol, S., and Melton, D.A. (1990). A two-step model for the localization of maternal mRNA in *Xenopus* oocytes: involvement of microtubules and microfilaments in the translocation and anchoring of Vg1 mRNA. *Development* *108*, 289-298.
- Zalfa, F., Adinolfi, S., Napoli, I., Kuhn-Holsken, E., Urlaub, H., Achsel, T., Pastore, A., and Bagni, C. (2005). Fragile X mental retardation protein (FMRP) binds specifically to the brain cytoplasmic RNAs BC1/BC200 via a novel RNA-binding motif. *J Biol Chem* *280*, 33403-33410.
- Zamore, P.D., Williamson, J.R., and Lehmann, R. (1997). The Pumilio protein binds RNA through a conserved domain that defines a new class of RNA-binding proteins. *RNA* *3*, 1421-1433.
- Zhang, H.L., Eom, T., Oleynikov, Y., Shenoy, S.M., Liebelt, D.A., Dichtenberg, J.B., Singer, R.H., and Bassell, G.J. (2001). Neurotrophin-induced transport of a beta-actin mRNP complex increases beta-actin levels and stimulates growth cone motility. *Neuron* *31*, 261-275.

REFERENCES

- Zhou, Y., and King, M.L. (1996). RNA transport to the vegetal cortex of *Xenopus* oocytes. *Dev Biol* *179*, 173-183.
- Zimyanin, V.L., Belaya, K., Pecreaux, J., Gilchrist, M.J., Clark, A., Davis, I., and St Johnston, D. (2008). In vivo imaging of oskar mRNA transport reveals the mechanism of posterior localization. *Cell* *134*, 843-853.

6 Appendix

6.1 Secondary-structure predictions of zipcode elements and of unrelated stem-loop containing RNAs

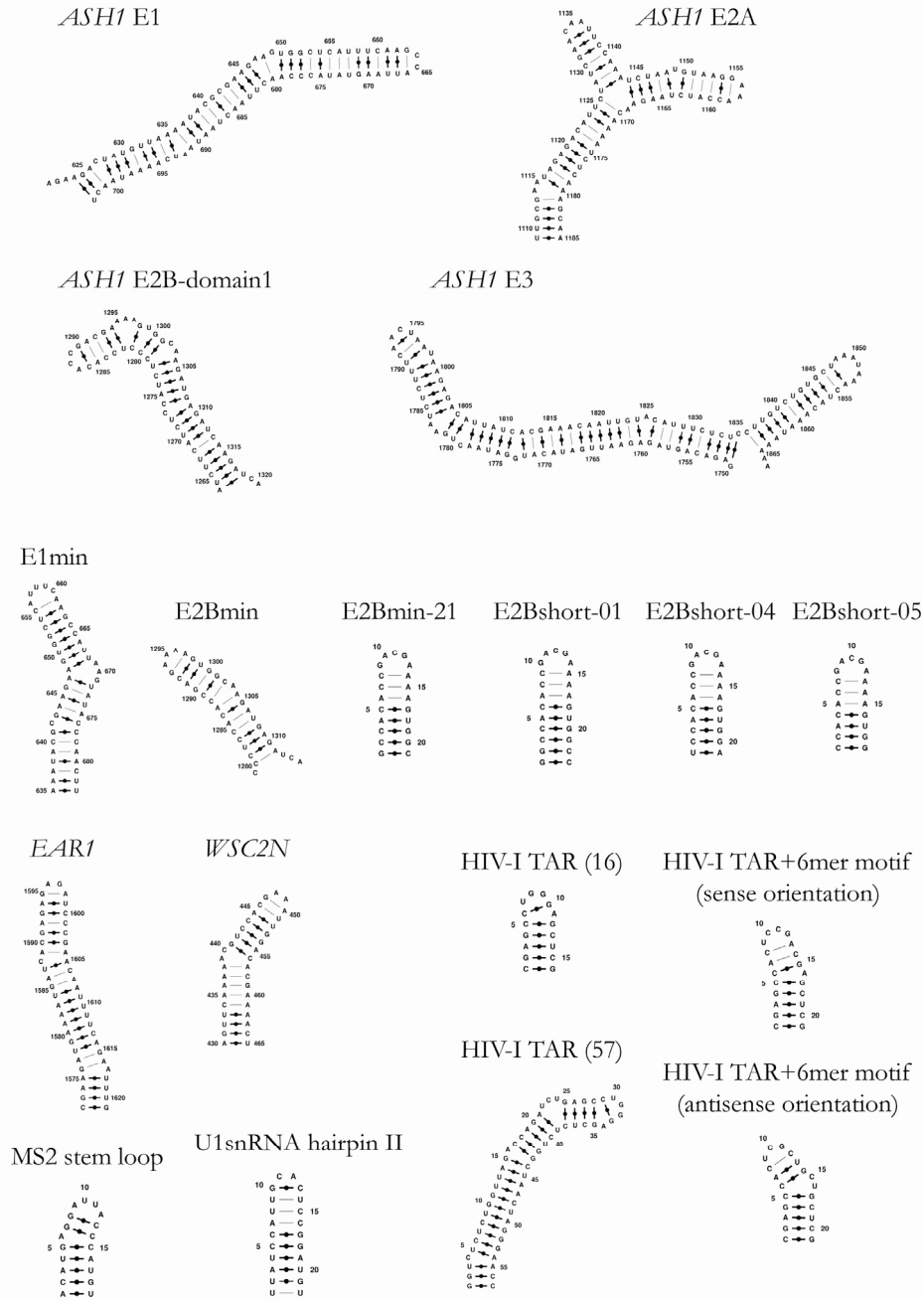


Figure A1: Secondary-structure predictions of zipcode elements and of unrelated stem-loop containing RNAs. Illustrated are the four *ASH1* zipcodes, the minimized *ASH1* E1 and *ASH1* E2B zipcodes used for crystallization, the zipcode of *EAR1*, the N-terminal zipcode of *WSC2*, and the stem loops of HIV-I TAR, U1snRNA, and MS2, respectively. Secondary predictions were calculated using the MC-Fold algorithm (Parisien and Major, 2008).

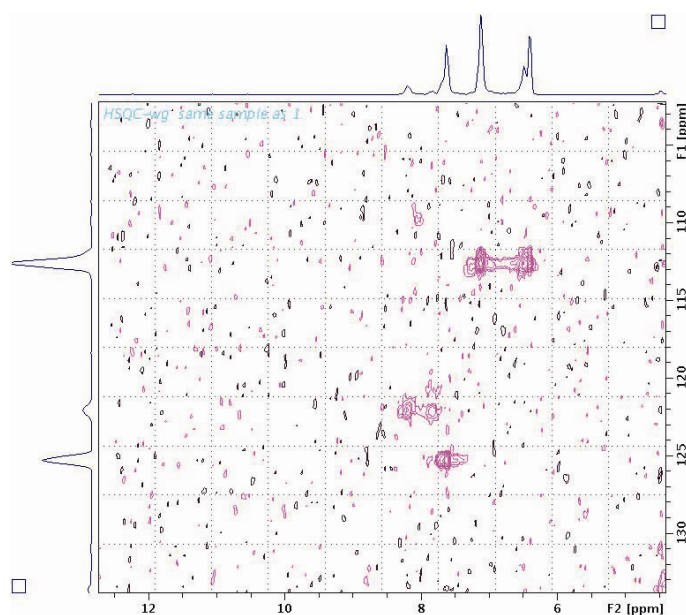
6.2 RNA oligonucleotides used for co-crystallization experiments with She2p-(6-246)- Δ Cys

Table A1: RNA oligonucleotides used for co-crystallization experiments with She2p-(6-246)- Δ Cys. The table summarizes the crystallization experiments that were performed with She2p-(6-246)- Δ Cys and the respective *ASH1*-mRNA fragments. The labels “+” and “-” indicate if crystals of the respective RNA:protein complex could be obtained or not. The sequences of all RNAs are listed in chapter 4.2.2.

RNA	length (nts)	Crystals in complex with She2p-(6-246)- Δ Cys
<i>ASH1</i> E3	118	-
E1min	49	-
E2Bmin	36	-
E2Bmin-21	21	+
E2Bshort-01	23	+
E2Bshort-02	24	-
E2Bshort-03	24	-
E2Bshort-04	21	+
E2Bshort-05	19	+
E2Bshort-06	20	-
E2Bshort-07	21	-
E2Bshort-08	19	-
E2Bshort-09	19	-

6.3 2D ^1H - ^{15}N -HSQC spectrum of wild-type She2p

Figure A2: 2D ^1H - ^{15}N -HSQC spectrum of wild-type She2p. Shown is a two-dimensional HSQC spectrum of She2p, which was recorded by the laboratory of Prof. Michael Sattler. Only a few peaks of She2p were obtained in the NMR experiments, rendering She2p inappropriate for NMR-titration experiments (see chapter 2.1.7).



6.4 She2p-sequence alignment

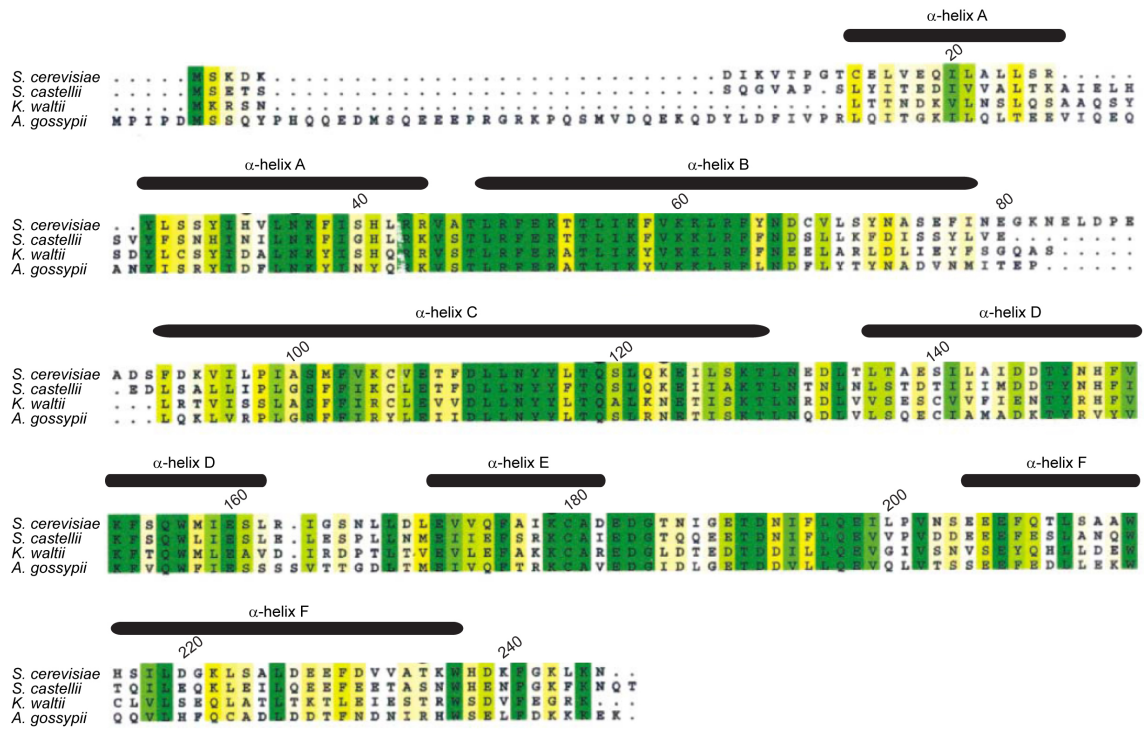


Figure A3: She2p-sequence alignment. Shown is a sequence alignment of She2p from different yeast species. Secondary structural elements that were obtained from the X-ray structure are denoted. Sequence similarity is encoded by a yellow-to-green color gradient, representing 40% to 100% identity. Figure was adapted from Niessing et al., 2004.

6.5 Assessment of *ASH1*-mRNA localization by *in situ* hybridization

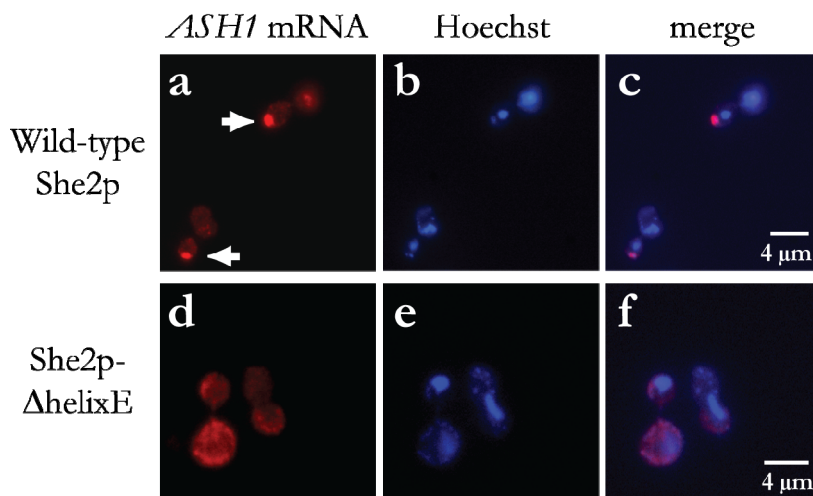


Figure A4: Impaired *ASH1*-mRNA binding by She2p results in mislocalization of *ASH1* mRNA *in vivo*. *In situ* hybridization experiments using fluorescently labeled oligonucleotides against *ASH1* were performed in wild-type cells and in She2p-ΔhelixE-expressing yeast strains. **a)-c)** In wild-type cells *ASH1* mRNA efficiently localizes to the bud tip (marked by an arrow). **d)-f)** In She2p-ΔhelixE-expressing $\Delta she2$ cells no *ASH1*-mRNA localization to the bud tip was observed (0 % localization in $n \geq 100$ cells). **(a, d)** show *in situ* hybridization with Texas Red-labeled oligonucleotides against *ASH1* mRNA, **(b, e)** show Hoechst nuclear staining, and **(c, f)** display merged signals of the respective *ASH1* and nuclei stainings. This experiment was performed by Maria Schmid from the laboratory of Prof. Ralf-Peter Jansen.

6.6 Sedimentation profile of wild-type She2p as obtained in sedimentation-velocity experiments

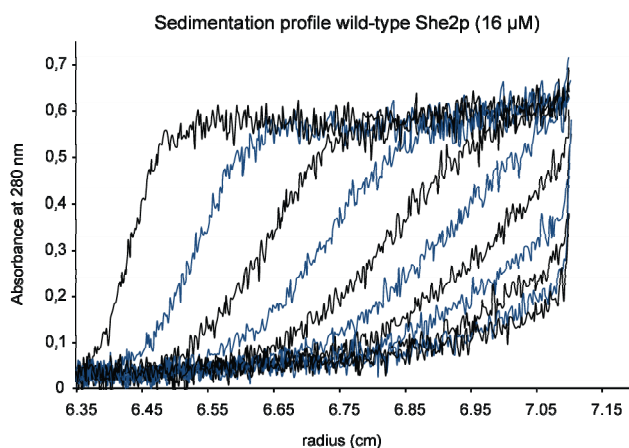


Figure A5: Sedimentation profile of wild-type She2p from a sedimentation-velocity experiment. Shown is the protein distribution during a sedimentation-velocity experiment, which was performed with wild-type She2p at an intermediate concentration of 16 μM. The absorbance measured at 280 nm is plotted against the radius of the measuring cell. The curves from left to right represent the protein distribution at increasing time points. From this profile, a sedimentation coefficient of 9.64 Svedberg (S) was calculated.

Curriculum Vitae

Marisa Müller

Personal Data

Date of birth: December 3, 1980
 Place of birth: Pforzheim, Germany
 Citizenship: German
 Marital status: Single

Education

1987 – 1991 Weiherbergschule Pforzheim – Grundschule
 1991 – 2000 Reuchlin Gymnasium Pforzheim
 Abitur

Academic Training

2000 – 2005 Biology studies at the Friedrich-Alexander University, Erlangen
 Diploma degree in Biology
 2004 – 2005 Diploma thesis at the Gene Center of the Ludwig-Maximilians-
 University Munich in the laboratory of Prof. Dr. Karl-Peter Hopfner
 “Struktur-Funktions-Untersuchungen von DNA-Reparatur
 ATPasen”
 2005 – 2009 Graduate studies at the Gene Center of the Ludwig-Maximilians-
 University Munich in the laboratory of Dr. Dierk Niessing
 “Characterization of She2p-dependent mRNP assembly in
 Saccharomyces cerevisiae”
 2006 – 2008 PhD fellowship from the Boehringer Ingelheim Fonds

Publications

Parts of this work have been published or are in process of publication:

Müller, M., Richter, K., Heuck, A., Kremmer, E., Buchner, J., Jansen, R.P., Niessing, D. (2009)
Formation of She2p tetramers is required for mRNA binding, mRNP assembly, and localization

Submitted

Du, T.G., Jellbauer, S., Müller, M., Schmid, M., Niessing, D., and Jansen, R.P. (2008)
Nuclear transit of the RNA-binding protein She2 is required for translational control of localized *ASH1* mRNA
EMBO Rep 9, 781-787

Other publications:

Heuck, A., Du, T.G., Jellbauer, S., Richter, K., Kruse, C., Jaklin, S., Müller, M., Buchner, J., Jansen, R.P., and Niessing, D. (2007)
Monomeric myosin V uses two binding regions for the assembly of stable translocation complexes
Proc Natl Acad Sci U S A 104, 19778-19783

Müller, M., Heuck, A., and Niessing, D. (2007)
Directional mRNA transport in eukaryotes: lessons from yeast
Cell Mol Life Sci 64, 171-180

Rigali, S., Nothhaft, H., Noens, E.E., Schlicht, M., Colson, S., Müller, M., Joris, B., Koerten, H.K., Hopwood, D.A., Titgemeyer, F., van Wezel, G.P. (2006)
The sugar phosphotransferase system of *Streptomyces coelicolor* is regulated by the GntR-family regulator DasR and links N-acetylglucosamine metabolism to the control of development
Mol Microbiol 61, 1237-1251

Dürr, H., Körner, C., Müller, M., Hickmann, V., and Hopfner, K.P. (2005)
X-ray structures of the *Sulfolobus solfataricus* SWI2/SNF2 ATPase core and its complex with DNA
Cell 121, 363-373

Acknowledgements

I would like to thank my supervisor Dierk Niessing for giving me the opportunity to work in his laboratory in an excellent environment at the Gene Center. I am grateful for his guidance over the years and especially for allowing me to pursue my own strategies.

I want to thank the members of my thesis advisory committee Patrick Cramer, Klaus Förstemann, and Ralf-Peter Jansen for supporting my work over the time in a very constructive way. Especially, I am grateful to Patrick and Klaus for being the first and second examiners of my thesis.

I am grateful to Ralf-Peter Jansen for his accordance to collaborate with us on my project. He was always open for exciting discussions about the She2p mystery and I could benefit a lot from his deep knowledge on mRNA localization in yeast.

I would like to thank all members of the Niessing group for their contribution to have a really nice time in the past four years. Especially, I want to thank ... Sigrun for being the “real” lab head keeping the lab running and for her great support ... Alexander for being always ready for discussions on our projects, even if there was nothing to discuss ;-) ... Daniela for critical comments on this thesis.

I want to thank Maria Schmid, Stephan Jellbauer, and Susanne Lange from the Jansen lab for their patience to introduce me into the work with yeast. I want to thank Klaus Richter and Johannes Buchner for performing the analytical ultracentrifugation experiments. I am very grateful to Sophia Hartung and Gregor Witte for introducing me into the SAXS technique. It was a pleasure to work with you! Alex Schele, Sophia Hartung, and Lina Karakasili I want to thank not only for having a great time together at the Gene Center, but also for nice activities outside the lab.

I am grateful to the Boehringer Ingelheim Fonds for financial support, but also for the nice atmosphere during the regular meetings and the continuous personal support. I am very glad for this fantastic experience!

Finally, and most importantly, I am deeply grateful for having a great and understanding family, who generously supported me over the last years. Thank you, Andi, for your constant support and for always being there for me!